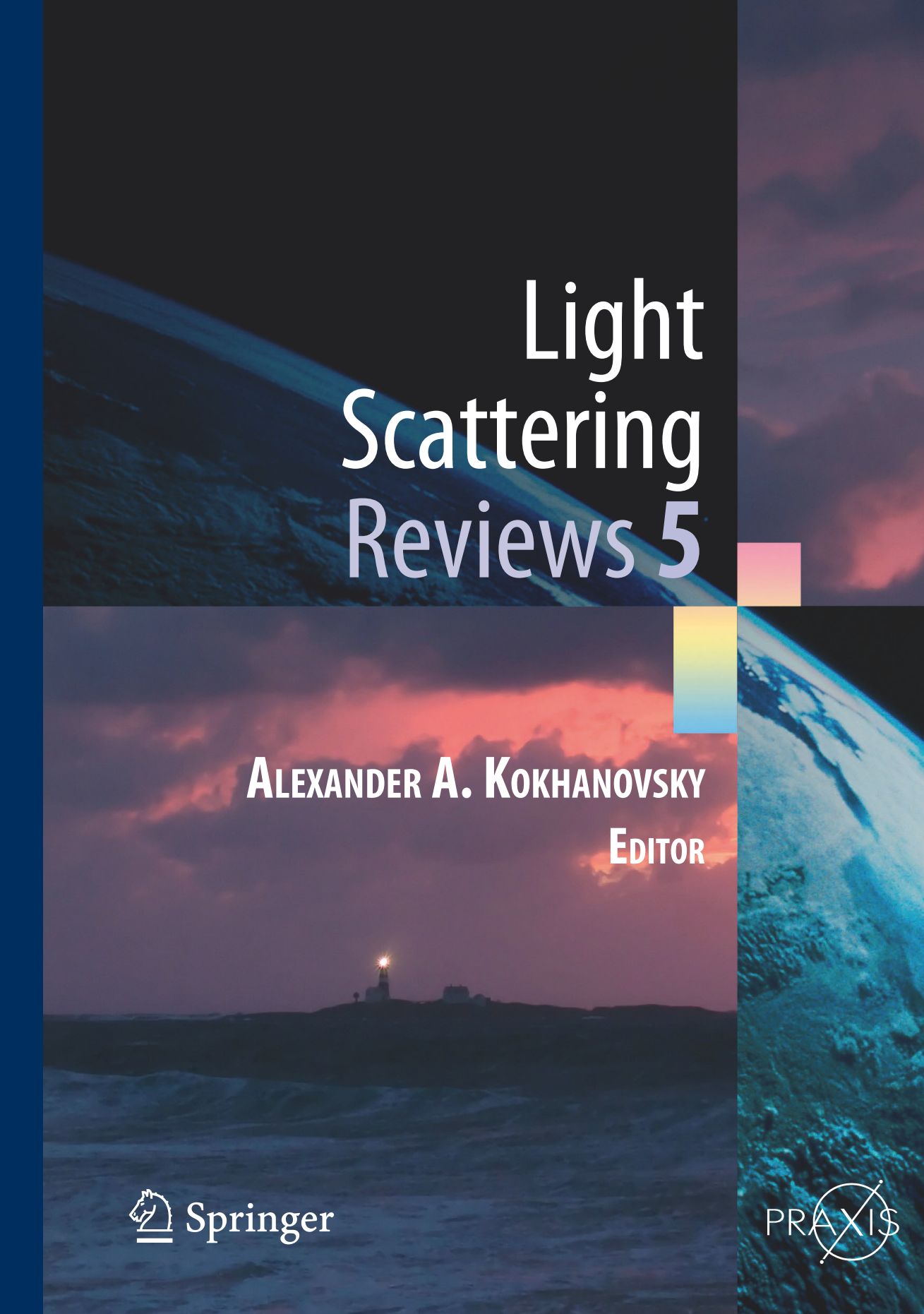

Light
Scattering
Reviews 5
Alexander A. Kokhanovsky
Editor
Springer
PRAXIS

Light Scattering Reviews 5

Single Light Scattering and Radiative Transfer

Alexander A. Kokhanovsky (Editor)

Light Scattering Reviews 5

Single Light Scattering and Radiative Transfer

Springer

Published in association with
Praxis Publishing
Chichester, UK

PRAXIS

Editor
Dr Alexander A. Kokhanovsky
Institute of Environmental Physics
University of Bremen
Bremen
Germany

SPRINGER–PRAXIS BOOKS IN ENVIRONMENTAL SCIENCES (*LIGHT SCATTERING SUB-SERIES*)
SUBJECT *ADVISORY EDITOR*: John Mason M.B.E., B.Sc., M.Sc., Ph.D.
EDITORIAL *ADVISORY BOARD MEMBER*: Dr Alexander A. Kokhanovsky, Ph.D. Institute of Environmental Physics, University of Bremen, Bremen, Germany

ISBN 978-3-642-10335-3 Springer Berlin Heidelberg New York

Cover design: Jim Wilkie
Project copy editor: Mike Shardlow
Author-generated LaTex, processed by EDV-Beratung Herweg, Germany

Printed in Germany on acid-free paper

Springer is part of Springer-Science + Business Media (springer.com)

Contents

List of Contributors

Vladimir P. Budak
Moscow Power Engineering Institute
Krasnokazarmennaya, 14
Moscow 111250
Russia
BudakVP@mpei.ru

Celine Cornet
Laboratoire d'Optique Atmospherique
UMR CNRS 8518
Université des Sciences et Technologies de Lille
Villeneuve d'Ascq
France
cornet@loa.univ-lille1.fr

Per Edström
Department of Natural Sciences, Engineering and Mathematics
Mid Sweden University
SE-871 88 Härnösand
Sweden
per.edstrom@miun.se

Sandra Jacquier
Ecole Nationale Supérieure des Mines
158 Cours Fauriel
42023 Saint-Etienne
France
jacquier_sj@yahoo.fr

Frédéric Gruy
Ecole Nationale Supérieure des Mines
158 Cours Fauriel
42023 Saint-Etienne
France
gruy@emse.fr

Teemu Hakala
Finnish Geodetic Institute
PO Box 15
02431 Masala
Finland
teemu.hakala@fgi.fi

Manuela Hirschmugl
Joanneum Research – Institut für Digitale Bildverarbeitung
Wastiangasse 6
8010 Graz
Austria
manuela.hirschmugl@joanneum.at

Sanna Kaasalainen
Finnish Geodetic Institute
PO Box 15
02431 Masala
Finland
Sanna.Kaasalainen@fgi.fi

Dmitriy A. Klyuykov
Moscow Power Engineering Institute
Krasnokazarmennaya, 14
Moscow 111250
Russia
klyuykovda@gmail.com

Sergey V. Korkin
Moscow Power Engineering Institute
Krasnokazarmennaya, 14
Moscow 111250
Russia
korkins@gmail.com

Laurent C-Labonnote
Laboratoire d'Optique Atmosphérique
UMR CNRS 8518
Université de Lille 1
59655 Villeneuve d'Ascq
France
Laurent.Labonnote@univ-lille1.fr

Jochen Landgraf
Netherlands Institute for Space Research (SRON)
Sorbonnelaan 2
3584 CA Utrecht
The Netherlands
J.Landgraf@sron.nl

Bonnie Light
Polar Science Center
Applied Physics Laboratory
University of Washington, Seattle
WA 98105
USA
bonnie@apl.washington.edu

Alexei Lyapustin
NASA Goddard Space Flight Center
Mail code 614.4
Greenbelt, MD 20771
USA
Alexei.I.Lyapustin@nasa.gov

Karri Muinonen
Department of Physics, PO Box 64,
Gustaf Hällströmin katu 2a
FI-00014 University of Helsinki
Finland
muinonen@cc.helsinki.fi

Tolegen Muldashev
101 Miras microdistrict
Almaty 050000
Kazakhstan
tmuldashev@hotmail.com

Jyri Näränen
Finnish Geodetic Institute
PO Box 15
02431 Masala
Finland
jyri.naranen@fgi.fi

Yasuhiko Okada
Nishi 1-729, Nakayama-Cho
Nara-City
Nara 631-0013
Japan
yasokada@gmail.com

Jouni Peltoniemi
Finnish Geodetic Institute
PO Box 15
02431 Masala
Finland
jouni.peltoniemi@fgi.fi

Eetu Puttonen
Finnish Geodetic Institute
PO Box 15
02431 Masala
Finland
eetu.puttonen@fgi.fi

Nikolai N. Rogovtsov
Applied Mathematics Department
Belarusian National Technical University
Prospect Nezavisimosty 65
220013 Minsk
Belarus
rogovtsov@bntu.by

F. Szczap
Laboratoire de Météorologie Physique
UMR CNRS 6016
Université Blaise Pascal
63177 Aubière
France
szczap@opgc.univ-bpclermont.fr

Juha Suomalainen
Finnish Geodetic Institute
PO Box 15
02431 Masala
Finland
juha.suomalainen@fgi.fi

Johanna Torppa
Observatory
PO Box 14
00014 University of Helsinki
Finland
johanna.torppa@helsinki.fi

Jani Tyynelä
Department of Physics
PO Box 68
FI-00014 University of Helsinki
Finland
jani.tyynela@helsinki.fi

Gorden Videen
Army Research Laboratory
2800 Powder Mill Rd.
Adelphi, Maryland 20783
USA
gorden.videen@arl.army.mil

Holger H. Walter
Albstrasse 29
73457 Essingen-Lauterburg
Germany
dr.holger.walter@googlemail.com

Yujie Wang
NASA Goddard Space Flight Center
Mail code 614.4
Greenbelt, MD 20771
USA.
Yujie.Wang@nasa.gov

Evgenij Zubko
Observatory
PO Box 14
Kopernikuksentie 1
FI-00014 University of Helsinki, Helsinki
Finland
evgenij.zubko@helsinki.fi

Notes on the contributors

Vladimir P. Budak is a professor at the MPEI (TU) (Moscow Power-Engineering Institute, Technical University), Light Engineering Department. His research activities are mainly devoted to the theory of radiative transfer, the development of the numerical solution method of the radiative transfer equation, and mathematical modeling of the signals in remote sensing systems. He received his PhD in 1985 and Doctor of Science degree in 1998 in optics from MPEI (TU). He has published over sixty refereed papers, books, and chapters in edited volumes, where the most recent ones are focused on polarized radiative transfer, 3-D radiative transfer, and image transfer through a turbid medium slab.

Céline Cornet received her MS degree in 2000 and her PhD in Atmospheric Sciences in 2003, both at the University Blaise Pascal of Clermont-Ferrand, France. She is currently assistant professor in Physics at the University of Lille 1, France, and scientist in the Laboratoire d'Optique Atmosphérique of Lille, France. Her research interests concern atmospheric radiative transfer modeling and cloud remote sensing. She is working more particularly on the effects of three-dimensional cloud structures for both direct radiative calculations and cloud parameters retrieval.

Per Edström graduated in Engineering Physics at Uppsala University, Sweden, in 1995, with a double major in Electrophysics and Systems Engineering. He received his PhD in Mathematics at Mid Sweden University, Sweden, in 2007, with a thesis titled *Mathematical Modeling and Numerical Tools for Simulation and Design of Light Scattering in Paper and Print*. His primary research interest is in applied mathematics, and he has been investigating properties of the forward and inverse radiative transfer problems, as well as developing and studying numerical methods for these. One important application area is industrial use in the paper and printing area, and he is currently national coordinator of paper optics research in Sweden. Dr Edström is a member of the Society of Industrial and Applied Mathematics.

Frédéric Gruy is a professor at the Ecole Nationale Supérieure des Mines, Saint-Etienne, France. He graduated from Ecole Normale Supérieure, Cachan, France, in 1979. He received the MS degree from Paris XI University, Orsay, France, where he studied at the department of physical chemistry. He obtained his PhD at ENSM on Saint-Etienne, France. His PhD work was devoted to grain coarsening in gas–solid systems. Since 1986 he has worked at ENSM topics related to the physics and physical chemistry of powders or particle suspensions. His main interest was the search for mechanisms underlying the change of the particle population with time. Collective behaviors of particles in suspension, particularly aggregation and agglomeration, were at the centre of his research. He has also worked on suspension characterization by optical methods since 1991. His current research is focused on the study of the relationship between geometry–topology and the optical or mechanical properties of particles and aggregates. His work is connected to a wide field of applications including ceramics, nano- and micro-particle suspensions and multiphase flow.

Sandra Jacquier obtained her PhD from the Ecole Nationale Supérieure des Mines de Saint-Etienne in 2006. Her thesis work addressed the search for theoretical techniques to study the optical properties of non-absorbent, aggregated, spherical particles. Since then, her scientific interests have dealed with the use and modeling of light–matter interaction to characterize or control materials or processes.

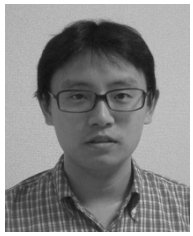
Jochen Landgraf is a senior research scientist at the Netherlands Institute for Space Research (SRON) and his main research fields are radiative transfer, trace gas retrieval, and photochemistry modeling. He obtained his PhD at the Max-Planck Institute for Chemistry in Mainz, Germany, where he developed a module for efficient online calculations for photodissociation rates. In addition, he has been involved in studying the effects of 3-D radiative transport in inhomogeneous cloud fields and their relevancy for global chemistry modeling. At SRON his research is focused on remote sensing from satellite measurements, particularly from GOME, GOME-2, and SCIAMACHY observations. His recent work has concerned new remote sensing concepts to retrieve tropospheric ozone, carbon monoxide, and methane using satellite measurements in the UV, visible and NIR in combination with emission spectra in the thermal infrared.

Bonnie Light is a research scientist at the Polar Science Center at the Applied Physics Laboratory of the University of Washington in Seattle. She has made both experimental and theoretical contributions to our understanding of how sea ice reflects, transmits, and absorbs sunlight. Her work involves laboratory studies, modeling, and field work, including four months spent at ice station SHEBA in the Beaufort Sea in 1998. Her interests in radiative transfer and sea ice have grown from a background in Electrical Engineering,

a passion for the environment, and a fascination with the Arctic. She received her PhD from the Department of Atmospheric Sciences at the University of Washington in 2000.

Alexei I. Lyapustin received an MS/BS in physics from Moscow State University in 1987, and PhD in Aerospace Remote Sensing from the Space Research Institute, Moscow, Russia in 1991. He is an associate research scientist at the University of Maryland Baltimore County. A. Lyapustin actively works in 1-D and 3-D radiative transfer in planetary atmospheres. His current research interests are focused on retrievals of atmospheric aerosol and surface bidirectional reflectance from MODIS and future NPOESS VIIRS sensors using time series analysis. Alexei is a science team member of the MODIS and NPP Science Teams, and of the GOES-R Algorithm Working Group, and is an associate member of the VIIRS Operational Algorithm Team (VOAT).

Karri Muinonen is a professor at the University of Helsinki and at the Finnish Geodetic Institute. His research entails light scattering by small solar-system particles, including both single scattering by individual particles and multiple scattering by complex random media of particles. He studies orbit determination for asteroids, particularly near-Earth objects, as well as the physical properties of small solar-system objects. He received his PhD in Theoretical Physics in 1990 and his Docentship in Space Astronomy in 1994, both at the University of Helsinki. He has worked as NASA Postdoctoral Fellow at Lowell Observatory (1990–1993) and Academy Research Fellow (1998–2003) at the University of Helsinki with an extended visit to Osservatorio Astronomico di Torino (2001–2002). He has further paid extended visits to Dipartimento di Matematica, Universita di Pisa (1996), and Astronomiska Observatoriet, Uppsala Universitet (1997). He has published over one hundred papers in peer-reviewed scientific journals.



Yasuhiko Okada graduated from the Interdisciplinary Graduate School of Science and Engineering, Kinki University, Osaka, Japan, in March 2003. He received his PhD (in Engineering) in the field of remote sensing of atmospheric aerosols, especially focusing on aerosol retrieval over turbid water. In October 2003, he joined "The 21st Century COE Program of Origin and Evolution of Planetary Systems", Graduate School of Science and Technology, Kobe University, Japan. Throughout the COE Program (until March 2008), he focused on research into light scattering by non-spherical particles based on numerical light scattering simulation methods. During the period between April 2008 and March 2009, he joined the Interdisciplinary Graduate School of Science and Engineering, Kinki University, Osaka, Japan. Since April 2009, he has started to work for a company as an engineer. He is a member of the American Geophysical Union, the Optical Society of America, and the Japanese Society for Planetary Sciences.

Jouni Peltoniemi received the PhD degree in theoretical physics from the University of Helsinki in 1993. He is currently a researcher at the Finnish Geodetic Institute and an adjunct professor at the University of Helsinki. His research interests are measurements, modeling and applications of light scattering from land surfaces and atmospheres. He has worked as research assistant at the University of Helsinki (1987–1993), as a post-doctoral researcher at Jena, Group "Dust in star forming regions" (1993–1995), and as a research scientist at the Finnish Geodetic Institute since 1995. His has led the development of several measurement instruments, including the Finnish Geodetic Institute Field Goniospectropolarimeter (FIGIFIGO) and UAV-based observation systems.

Nikolai N. Rogovtsov graduated from the Physical Department of the Belarusian State University, Minsk, Belarus, in 1970. He received his PhD in optics and his Doctor of Sciences degree in physics and mathematics from the B.I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Belarus, in 1976 and 1994, respectively. He is currently a professor of Mathematics at the Applied Mathematics Department of the Belarusian National Technical University. His research interests cover the areas of radiative transfer theory (RTT), invariance theory, theory of infinite linear systems of algebraic and differential equations, and theory of integral and integro-differential equations. In 1981 he formulated a general invariance principle, giving his algebraic and physical interpretation. During 1976–1983 he developed a general invariance relations reduction method. In 1996 Professor Rogovtsov solved in an analytical form the characteristic equation of RTT. Professor Rogovtsov is the author of the monograph "Properties and principles of invariance. Application to Problems of Mathematical Physics" (1999, in Russian). He is the author and co-author of more than 100 papers.

Holger Walter studied meteorology at the Johannes-Gutenberg University in Mainz, Germany (1995–2000). He specialized in the transport of radiation in the terrestrial atmosphere. Afterwards he was employed at the Netherlands Institute for Space Research (SRON) in Utrecht, The Netherlands. There he developed a linearized radiative transfer model for the interpretation of satellite measurements in limb viewing geometry. In 2006 he received a PhD at the Vrije Universiteit of Amsterdam, The Netherlands. Since 2006 he has been a scientist at Carl Zeiss SMT AG, Oberkochen, Germany. He currently works on optical problems in the domain of microlithography.

Preface

Light scattering by densely packed inhomogeneous media is a particularly challenging optics problem. In most cases, only approximate methods are used for the calculations. However, in the case where only a small number of macroscopic scattering particles are in contact (clusters or aggregates) it is possible to obtain exact results solving Maxwell's equations. Simulations are possible, however, only for a relatively small number of particles, especially if their sizes are larger than the wavelength of incident light. The first review chapter in Part I of this volume, prepared by Yasuhiko Okada, presents modern numerical techniques used for the simulation of optical characteristics of densely packed groups of spherical particles. In this case, Mie theory cannot provide accurate results because particles are located in the near field of each other and strongly interact. As a matter of fact, Maxwell's equations must be solved not for each particle separately but for the ensemble as a whole in this case. The author describes techniques for the generation of shapes of aggregates. The orientation averaging is performed by a numerical integration with respect to Euler angles. The numerical aspects of various techniques such as the T-matrix method, discrete dipole approximation, the finite difference time domain method, effective medium theory, and generalized multi-particle Mie solution are presented. Recent advances in numerical techniques such as the grouping and adding method and also numerical orientation averaging using a Monte Carlo method are discussed in great depth.

The second chapter in Part I of this volume, prepared by Sandra Jacquier and Frédéric Gruy, also considers the scattering properties of aggregates. The authors consider particles with radii both much smaller and larger as compared to the wavelength of the incident light. In the least case the methods based on the direct solution of Maxwell's equations cannot be used due to computational problems. Therefore, various approximate techniques are introduced and thoroughly discussed in the chapter. They include the anomalous diffraction approximation developed by van de Hulst, the Berry–Percival–Khlebtsov method, effective medium technique, and a compact sphere method. The performance of these methods (where it is possible) is evaluated against exact computations. As an application of theoretical techniques, the authors consider the process of agglomeration of small primary particles in a homogeneous suspension and its monitoring using measurements of turbidity.

Part II of this book deals with radiative transfer theory. This theory is used to describe reflectance and transmittance of turbid media such as the atmosphere and the ocean. It is based on solution of the vector radiative transfer equation (VRTE). Usually the VRTE is solved using the approximation of horizontally homogeneous

media. However, such an approach cannot be used in many cases (e.g., in the vicinity of cloud edges) and the 3-D VRTE must be formulated and solved. Céline Cornet and co-authors describe one possible way of solving the 3-D VRTE. It is based on the Monte Carlo approach. The technique implemented in the developed software (code 3DMCpol) is described together with validation using the adding-doubling method and also the Spherical Harmonics Discrete Ordinate Method (SHDOM). The code is applied to study the Stokes vector of reflected light for a synthetic heterogeneous cloud field.

The authors of the second chapter in Part II, Holger Walter and Jochen Landgraf, deal with 3-D radiative transfer as well. However, they ignore polarization characteristics of light and concentrate their efforts on the solution of the radiative transfer equation in a spherical geometry, which is of importance for atmospheric remote sensing, including satellite limb measurements. The Picard iteration method is used with validation from reference Monte Carlo calculations. Also, the authors perform the linearization of the radiative transfer problem in the spherical geometry. The forward and adjoint radiative transfer equations (ARTE) are introduced. ARTE differs from the standard radiative transfer equation by the sign of the streaming term d/ds (also the directions in the scattering function are reversed). An important point is that ARTE can be solved using the same radiative transfer solver as applied for the solution of the standard forward radiative transfer equation (by exchanging and reversing the incoming and outgoing directions). The solutions of the forward and adjoint radiative transfer equations are used to estimate the impact of change in the atmospheric state on the light intensity in a given direction, which is of great importance both for sensitivity studies and for the inverse problem solution.

The next chapter (Chapter 5) prepared by Vladimir Budak and co-authors, describes a method for solving the radiative transfer problem with strongly anisotropic scattering (e.g., relevant to light propagation in the ocean, where the asymmetry parameter can reach 0.99). The subject is very important, and the authors offer an original and very effective solution based on a separation of regular (smooth) and highly anisotropic components. Comparison with a well-established discrete ordinate code (DISORT) is given. This confirms a high accuracy of the developed code (and, in addition, it has much higher speed as compared to DISORT in the case of the highly elongated in the forward direction phase functions). To eliminate the anisotropic part of the solution of the radiative transfer equation the authors used the small-angle approximation in Goudsmit–Saunderson form. It results in a boundary-value problem for the regular part of the solution that is similar to the initial problem, but with the modified source function on the right-hand side of equation and the modified boundary conditions. The solution of the obtained equation for the regular part is found by the discrete ordinates method. In the case of plane-parallel geometry this problem has the analytical solution as a matrix exponential. The authors also generalized their approach to account for the vector nature of light and solved VRTE using the same methodology. Moreover, the 3-D radiative transfer has also been considered in the framework of the generalized Goudsmit–Saunderson approximation.

Chapter 6, by Lyapustin and co-authors, provides an overview of the publicly available radiative transfer code SHARM. Among rigorous scalar codes, SHARM

is one of the most numerically efficient and is based on the method of spherical harmonics. Following descriptions of the algorithm and interface of the code, the chapter describes the code SHARM-3D, which is designed for computations with non-homogeneous anisotropic surfaces. SHARM-3D uses a rigorous 3-D Green's function solution for radiance in combination with the linear kernel model of the land surface bi-directional reflectance. In this case, the top-of-the-atmosphere radiance analytically depends on the surface BRDF parameters. This provides superior numerical efficiency, because when the atmosphere can be considered horizontally homogeneous, the atmospheric Green's function and the related functions need to be calculated only once. To handle the 'adjacency' effect, SHARM-3D uses a comprehensive pre-computed look-up table (LUT) of the 3-D optical transfer function (OTF) developed for different atmospheric conditions and geometries. OTF is a Fourier transform of the atmospheric 3-D Green's function. Certain symmetry properties of the OTF allow for a small LUT and a fast OTF-restore algorithm for arbitrary atmospheric conditions, geometry, and spatial resolution (currently $\geq$30 m). Comparisons of SHARM-3D with rigorous 3-D code SHDOM showed an agreement to 0.5–1%. At the same time, SHARM-3D is much faster (a factor of $\sim 10^3$) and, contrary to SHDOM, allows the processing of large scenes (side dimension of 10^3–10^4) on a common laptop.

As we see, all the chapters in Part II touch upon 3-D radiative transfer problems, which is a hot topic of modern radiative transfer research. Usually, light propagation in media with arbitrary spatial distributions of optical characteristics is studied using numerical calculations and it is a difficult matter to derive analytical results in this case. One possibility of the simplification of the 3-D solution is based on the principles of invariance. These principles were introduced in the radiative transfer theory by V. A. Ambartsumian, who derived the nonlinear integral radiative transfer equation for a semi-infinite layer using the fact that adding an additional layer at the top with the same properties as the medium itself will not change the reflected light intensity (by definition of a semi-infinite layer). Further, the principles of invariance have been explored by a number of prominent scientists such as S. Chandrasekhar and R. Bellman. In the concluding chapter of Part II, Nikolai Rogovtsov describes the physical and mathematical foundations of the general invariance relations reduction method (GIRRM). GIRRM is one of the most general methods of the solution for both direct and inverse radiative transfer problems. The method can be used as the basis for the development of both effective numerical schemes and analytical solutions of 3-D radiative transfer problems for arbitrary phase functions and spatial distributions in turbid media of various configurations. In addition, using the described approach, the author derived a new approximate representation for the reflection function of a semi-infinite plane-parallel medium. The asymptotic solutions valid if the characteristic dimensions of the disperse medium are much larger as compared to the photon mean path length in the medium are derived for a number of light scattering objects. In particular, disperse media with shapes of sphere, cube, cylinder, and spheroid have been considered. The derived results can be utilized, for example, for testing various numerical methods and Monte Carlo solutions of the 3-D problem.

The book concludes with Part III aimed at applications. Bonnie Light presents a comprehensive summary of sea ice optical properties and their relationships with

the structural properties of sea ice. Here the focus is not on light polarization and 3-D effects but rather on the quantification of the impact of various ice impurities such as brine inclusions, precipitated salt crystals, inclusions of biogenic and lithogenic origin, and bubbles on light scattering and radiative transfer in sea ice. The main problem here is the size and shape of inclusions, and also their refractive index. Inclusions can be several centimeters long and have irregular shapes. The sizes, shapes, and also refractive indices of some inclusions (e.g., brine) in ice strongly depend on temperature. Clearly, ice is a very difficult turbid medium to study in the field. Therefore, the author concentrated her study on the quantification of ice optical and microphysical properties in the cold laboratory, which resulted in the development of a physically based structural-optical model for first-year Arctic sea ice. The results of this comprehensive work are invaluable for the development of new methods of satellite remote sensing of ice, which are needed to quantify the effects of global change in the Arctic.

Ice is often covered by snow and, therefore, understanding snow optical properties in relation to the snow grain size, density, wetness, and level of pollution is also of great importance. This subject is explored by Jouni Peltoniemi and co-authors. Both theoretical methods (ray tracing) and comprehensive measurement field campaigns together with relevant instrumentation and data processing are described. The authors show not only the snow bi-directional reflection function but also the spectral and directional dependencies of the measured degree of polarization, which is low in the visible and increases dramatically in the vicinity of ice absorption bands in the near-infrared. The potential of optical methods for snow melt monitoring is discussed.

The next chapter, Chapter 10, prepared by Per Edström, is aimed at the modeling of light scattering in paper. Paper is close in appearance to snow and this is due to the fact that scattering elements are weakly absorbing and strong multiple scattering takes place. However, there are also some differences due to the size, shape, and refractive index of scattering elements in paper. The author describes various models used currently in the paper industry to understand and optimize paper appearance and spectral reflectance. They include Kubelka–Munk theory, the discrete ordinates method, the Monte Carlo technique, and asymptotic radiative transfer solutions.

The last chapter of this volume is also aimed at studies of reflection but for planetary regoliths. Their brightness is much smaller compared to snow, ice, and paper. As stated by Karri Muinonen and his co-authors, two ubiquitous phenomena are observed for the planetary regoliths near opposition (where the direction of illumination almost coincides with the viewing direction and backscattering effects dominate): negative linear polarization and nonlinear surge of brightness. The phenomena are observed at sun–object–observer angles of less than 30 and 10 degrees, respectively, sometimes showing up at extremely small phase angles within a degree from opposition. Coherent backscattering and shadowing mechanisms have been introduced to explain the phenomena. The authors have studied interference mechanisms in scattering by single particles capable of explaining such intensity and polarization phenomena. The presented modeling constitutes an important advance in the interpretation of the observations of atmosphereless solar-system

objects. As shown by the authors, their observed scattering characteristics can be linked to the physical characteristics of submicrometer-scale scatterers.

In conclusion, I thank all the contributors for the preparation of the excellent papers presented in this volume. I am also indebted to Clive Horwood, Publisher, for advice, patience, and encouragement.

Bremen, Germany
October, 2009

Alexander A. Kokhanovsky

Part I

Optical Properties of Small Particles and their Aggregates

1 Numerical simulations of light scattering and absorption characteristics of aggregates

Y. Okada

1.1 Introduction

A non-sphericity of a particle plays an important role in light scattering processes, resulting in different scattering and absorption of incident light compared with spherical particles. Among various non-spherical shapes of particles, aggregates of small particles are often applied as model shapes to particles observed actually in nature, such as dust from cometary nuclei, soot aerosols floating in the Earth's atmosphere, and microbiocells composed of sets of small organic cells.

The distinct difference between light scattering properties of aggregates and those of a single particle (e.g., compact spheres, spheroids, and hexagonal prisms) is that the aggregates multiply scatter incident light by composing particles. Because of this difference, aggregates and a single particle have totally different light scattering properties, which will cause largely different estimation of physical properties from the same observational data when the light scattering properties are used for the interpretation. Therefore it is important to investigate light scattering properties of aggregates when applied to particles likely having aggregated structures.

Light scattering properties can be investigated based on two methods: (1) laboratory measurements and (2) numerical light scattering simulations.

In the former method (i.e., laboratory measurements), aggregates with sizes much larger than actual ones are used for the study. Those large aggregates can be used to study the light scattering properties of small aggregates on the condition that the aggregate sizes compared with the observing wavelength are the same, that is, for the same size parameter $x = 2\pi r/\lambda$ where r is the radius of the aggregate, λ is the observing wavelength (see discussion about the scale invariance rule in [49]). Therefore, the small aggregates observed at the visible wavelength can be studied in laboratory measurements by investigating larger aggregates at the microwave wavelength (i.e., microwave analog method [27, 28]).

In contrast to laboratory measurement, in numerical light scattering simulations, the light scattering processes on aggregates are simulated on computers based on the computational solution methods of Maxwell's equations. Different solution methods for Maxwell's equations have been developed and used as shown later in section 1.3. Numerical light scattering simulations have the advantages of being

capable of treating homogeneity in the physical properties of the aggregate, such as the size, material composition, alignment of composing particles.

Despite the advantage of treating homogeneity in physical properties of the aggregate, numerical light scattering simulations still have limitations (e.g., the maximum treatable size, treatable complexity in the shape of the particle, longer computational time, and accuracy). Therefore, various methods are still being developed and improved in order to overcome these limitations so that numerical light scattering simulations will be fully applicable to the interpretation of actually observed aggregates, whose sizes and complex shapes are currently located outside the treatable region of the numerical light scattering simulations.

In this chapter, we will introduce various results of the numerical light scattering simulations, which are applicable to fractal aggregates. This chapter is composed as follows. In section 1.2, various parameters defined for numerical light scattering simulations are introduced. In section 1.3, numerical light scattering simulation methods proposed from the 1960s to the 1990s are briefly introduced. Then, methods widely used in recent light scattering studies are presented with the applicability to aggregates. In section 1.4, we introduce the improvements of the numerical light scattering simulations for the aggregates and a few sample results from the simulations. Section 1.5 summarizes this chapter.

1.2 Properties of aggregates used in numerical simulations

1.2.1 Physical and light scattering properties

We introduce parameters for physical properties of an aggregate and those for their light scattering properties obtained from the numerical light scattering simulations. Table 1.1 summarizes the physical properties of an aggregate and conditions used in the simulations.

Table 1.1. Physical properties of an aggregate and conditions of the simulations

Monomer radius:	r_m
Wavelength of incident light:	λ
Monomer size parameter:	$x_m = \dfrac{2\pi}{\lambda} r_m$
Number of monomers:	N
Volume equivalent size parameter of the aggregate:	$x_v = \sqrt[3]{N} x_m$
Complex refractive index:	$m = n + ik$
Shape of the aggregate:	BPCA and BCCA (see section 1.2.2)

The composing particles of an aggregate are called monomers. The size of the monomer is described by the size parameter of the monomer $x_m (= 2\pi r_m/\lambda)$ where r_m is the monomer radius, and λ the wavelength of incident light. r_m and λ are treated in the same unit (e.g., μm or nm).

The number of monomers composing the aggregate is defined with N. From x_m and N, we can calculate the volume equivalent size parameter of the aggregate x_v (see Table 1.1 for the equation).

The material composition of the aggregate is described by the complex refractive index $m = n + ik$. k is related to the absorption coefficients of the aggregate defined as $4\pi k/\lambda$ (see Eq. 2.69 in [31]). Throughout this chapter, we only treat an aggregate of a single material composition. However, it is also possible to treat a composite aggregate whose monomers have different material compositions.

From the light scattering simulations with the above physical properties of the aggregate, we can obtain the light scattering properties of the aggregate. As integrated light scattering properties, we have efficiencies for absorption (Q_{abs}), scattering (Q_{sca}), and extinction (Q_{ext}), where $Q_{ext} = Q_{sca} + Q_{abs}$. The efficiencies are equal to cross-sections (C_{abs}) divided by the geometrical cross-section (πr_v^2) where r_v denotes the radius calculated from x_v and λ.

In addition to the integrated light scattering properties, we also obtain angular light scattering properties called the phase matrix. The phase matrix of an aggregate averaged over random orientations is defined in the following equation for the Stokes vector $[I, Q, U, V]^T$ (see chapter 4 of [48]).

$$\begin{bmatrix} I^{sca} \\ Q^{sca} \\ U^{sca} \\ V^{sca} \end{bmatrix} \propto \begin{bmatrix} a_1(\Theta) & b_1(\Theta) & 0 & 0 \\ b_1(\Theta) & a_2(\Theta) & 0 & 0 \\ 0 & 0 & a_3(\Theta) & b_2(\Theta) \\ 0 & 0 & -b_2(\Theta) & a_4(\Theta) \end{bmatrix} \begin{bmatrix} I^{inc} \\ Q^{inc} \\ U^{inc} \\ V^{inc} \end{bmatrix}, \tag{1.1}$$

where Θ is the scattering angle, which is defined by directions of incidence and scattering of light. $\Theta = 0^\circ$ denotes forward scattering and $\Theta = 180^\circ$ backward scattering. *inc* and *sca* denote incident and scattered components of the Stokes vector, respectively.

The a_1 element of the phase matrix is called the phase function. a_1 describes the angular distribution of incident light energy when the normalization condition in the following equation holds:

$$\frac{1}{2}\int_0^\pi d\Theta \, \sin\Theta \, a_1(\Theta) = 1. \tag{1.2}$$

The degree of linear polarization $P(\Theta)$ for the unpolarized incident light (i.e., $Q^{inc} = U^{inc} = V^{inc} = 0$) is defined as follows (see p. 28–29 of [48]):

$$P(\Theta) = -b_1(\Theta)/a_1(\Theta). \tag{1.3}$$

Table 1.2. Parameters of the integrated light scattering properties

Extinction efficiency:	$Q_{ext} = C_{ext}/(\pi r_v^2)$
Scattering efficiency:	$Q_{sca} = C_{sca}/(\pi r_v^2)$
Absorption efficiency:	$Q_{abs} = C_{abs}/(\pi r_v^2)$
Cross-sections:	C_{ext}, C_{sca}, and C_{abs} have the unit of $(length)^2$
Asymmetry parameter:	$g = \langle\cos(\Theta)\rangle = \frac{1}{2}\int_{-1}^{1} d(\cos\Theta) a_1(\Theta) \cos\Theta$
Single scattering albedo:	$\omega_0 = Q_{sca} \ / \ Q_{ext}$

r_v is a radius of volume-equivalent-sphere calculated from x_v and λ

1.2.2 Shapes of aggregates

There are several types of aggregates with different alignments of monomers. We only consider aggregates composed of spheres. However, it is possible to consider aggregates composed of spheroids and other shapes in the numerical light scattering simulations (e.g., [82]).

Aggregates having different structures can be described using the fractal dimension D_f. $D_f = 2$ and 3 denote structures close to 2-D (plane) and 3-D (sphere), respectively. By investigating aggregates with these two extreme cases of D_f, the light scattering properties of aggregates having intermediate D_f can be constrained.

Aggregates having D_f of around 2 and 3 are prepared based on sequential ballistic collisions of monomers or of small aggregates. We use two models called BPCA (Ballistic Particle-Cluster Aggregate) and BCCA (Ballistic Cluster-Cluster Aggregate) [55].

The BPCA is created as follows. Firstly, one monomer collides with another monomer generating a bisphere. Then one monomer after another collides with the bisphere from random directions resulting in a single large aggregate.

The BCCA is created as follows:

- collision of two monomers,
- collision of two bispheres produced in the previous step,
- collision of two aggregates produced in the previous step,
- recursive collisions of two aggregates produced in the previous step until all the monomers are gathered into a single large aggregate.

The BPCA has D_f of 2.98 ± 0.02 while the BCCA has D_f of 1.93 ± 0.07 [55]. The porosities of the BPCA and BCCA are defined with characteristic radius $r_a = \sqrt{5/3} r_g$, where r_g denotes the radius of gyration, which is obtained from the following equation [56]:

$$ {r_g}^2 = \frac{1}{2N^2} \sum_{i,j}^{N} (\mathbf{r_i} - \mathbf{r_j})^2, \tag{1.4} $$

where $\mathbf{r_i}$ and $\mathbf{r_j}$ denote position vectors of the ith and jth monomers.

The porosity of the BPCA is around 85% while that of the BCCA is greater than 99% (see table 1 of [36]).

Fig. 1.1 shows examples of 3-D images of the BPCA and BCCA composed of 128, 256, 512, and 1024 monomers. Note that the structures of the aggregates become different for different generations of the aggregate (e.g., five generations for BPCA with $N = 128$). However, all the aggregates with different generations can be described with similar fractal dimension D_f, which result in similar light scattering properties between generations of a fixed shape model with the same N (e.g., BPCA512a, BPCA512b, and BPCA512c).

The BCCA stretches outward largely compared with the BPCA because of the larger porosities (e.g., $\geq$99%). The differences in porosities between the BPCA and BCCA influence the light scattering properties.

The BPCA and BCCA can commonly be treated in numerical light scattering simulations only by changing the file describing the positions of all the monomers (the shape file). Light scattering properties of aggregates having the same x_v, and

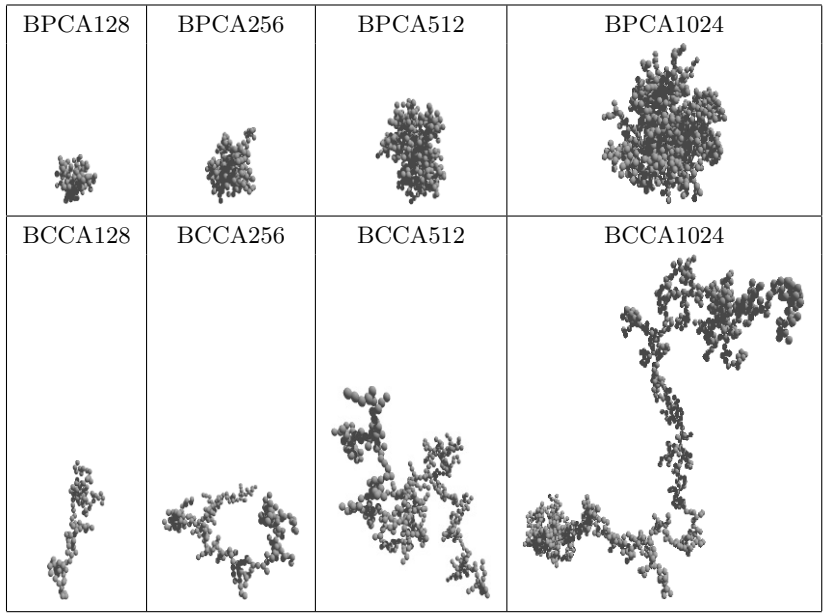


Fig. 1.1. Shape models; the BPCA and BCCA (after Y. Okada et al., *JQSRT*, **109**, 2613–2627, 2008).

the same m, but the different shapes of aggregates can be studied by changing the shape file.

1.2.3 Aggregate orientation

Light scattering properties are obtained for either

- a fixed orientation of the aggregate (a fixed direction of incident light),
- random orientation averaging.

The result of a fixed orientation is used to study an aggregate when the aggregate under observation is fixed with respect to the direction of incident light during the whole observation time. For example, fixed orientation results can be used to study the interstellar extinction of dust particles, which are possibly aligned by interstellar magnetic fields [32].

Results of random orientation averaging are used when the aggregate is rotated randomly during the observation time, or for an ensemble of particles whose monomers have similar shapes but have different orientations, seen in the observational field of view [47].

The orientation of the aggregate is defined with three Euler angles (i.e., α, β, and γ) with respect to three fundamental axes of a selected coordinate system [47] (see also Fig. 1.2).

Light scattering properties of random orientation averaging can be obtained by integrating results of fixed orientations (e.g., sets of (α, β, γ)). The values of the

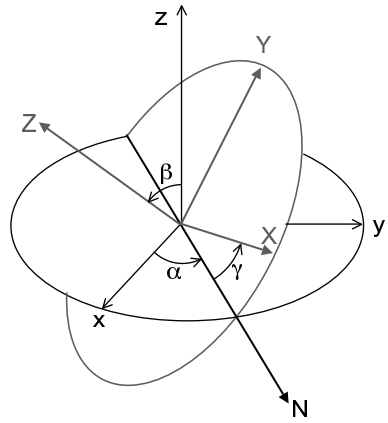


Fig. 1.2. Rotation of the aggregate using the Euler angles (α, β, γ). N denotes the line of node. Small capitals (x, y, z) denote orientation of the aggregate before the rotation. Capital letters (X, Y, Z) denote the orientation after the rotation.

random orientation averaging are calculated with the following equation:

$$\langle A \rangle = \frac{1}{8\pi^2} \int_0^{2\pi} \int_0^{\pi} \int_0^{2\pi} d\alpha \, \sin\beta \, d\beta \, d\gamma A(\alpha, \beta, \gamma), \tag{1.5}$$

where A denotes light scattering properties (e.g., cross-sections and phase matrix elements).

Random orientation averaging can be carried out based on two methods: (1) AOA (analytical orientation averaging), and (2) NOA (numerical orientation averaging). When the analytical solution for the orientation averaging is used, the random orientation averaging procedure can be carried out more efficiently. Therefore, deriving an analytical solution for the orientation averaging is an active research topic for the numerical light scattering simulations of irregularly shaped particles (see references in [51, 52, 86]).

The AOA is efficient for random orientation averaging. However, there are several numerical methods for light scattering simulations developed for a fixed orientation of the aggregate without the AOA solutions. When researchers use numerical methods of light scattering simulations developed only for a fixed orientation of the aggregate, they need to carry out the NOA to obtain results of random orientation averaging.

1.3 Methods for numerical light scattering simulations

We introduce numerical methods for simulating light scattering of aggregates. We focus on aggregates with $x_m \lesssim 2$ because aggregates with $x_m \gtrsim 2$ and $N \gtrsim 200$ are still currently difficult to be treated in rigorous numerical simulations.

Table 1.3 lists methods proposed by various researchers from the 1960s to the 1990s. Around the 1970s, numerical simulation methods were only applicable to

Table 1.3. Methods for multiple spheres proposed from the 1960s to the 1990s

- Liang and Lo (1967)
  - Proposed multipolar expansion method utilizing translational addition theorem for VSWF (vector spherical wave function)
  - Applicable to monomer radius less than $3\lambda/4$
- Bruning and Lo (1969), Bruning and Lo (1971a), and Bruning and Lo (1971b)
  - Derived recursion relation for translation coefficients in translational addition theorem
  - Applicable to monomer radius up to 10λ
- Jones (1979)
  - Derived a system of $3N$ linear equations for clusters of particles
  - Applicable to monomers in the Rayleigh region
- Borghese et al. (1979)
  - Proposed a method based on Debye potentials and mathematical technique to incorporate multiple scattering influence
  - Applicable to a system of molecules (radius less than 5 Å)
- Gérardy and Ausloos (1980)
  - Proposed a theory taking into account all multipolar orders
  - Applicable to N dielectric spheres embedded in dielectric matrix
- Gérardy and Ausloos (1982)
  - Incorporated higher-order multipolar interaction effects
  - Applicable to aggregated metal spheres
- Borghese et al. (1984a)
  - Derived a general addition theorem for vector Helmholtz harmonics
- Borghese et al. (1984b)
  - Utilized group theory for multipolar expansion of the light scattering simulations
- Fuller et al. (1986)
  - Incorporated an approach to treat non-interacting spheres
- Fuller and Kattawar (1988a), Fuller and Kattawar (1988b)
  - Proposed a method based on the OoS (order-of-scattering) technique
- Mackowski (1991)
  - Devised a new recurrence relation in the addition theorem for efficient implementation of the OoS technique
- Wang and Chew (1993)
  - Proposed RATMA (recursive aggregate T-matrix algorithm)
- Fuller (1994)
  - Proposed a method based on summation over pairwise cross-sections for calculations of cross-sections
- Borghese et al. (1992), Borghese et al. (1994), Fuller (1995)
  - Proposed a method for a sphere containing sub-sphere inclusions

simpler shapes, such as bispheres, and aggregates composed of up to five monomers, or the application was limited to aggregates composed of monomers whose sizes are in the Rayleigh region. Those methods were gradually improved for increased numbers of monomers and for increased sizes of monomers. Also, the computational efficiency of the numerical simulation method has been gradually improved.

Table 1.3 only lists the methods focusing on the treatment of multiple spheres. However, methods for shapes other than multiple spheres were also studied during the same period. For example, Mishchenko (1991) proposed a method for light scat-

tering by randomly oriented axially symmetric particles [46]. In his paper, rigorous analytical solution for the orientation averaging is introduced. The work improved the T-matrix method originally proposed in [75, 76] (see [51, 52, 86] and references therein).

Now we move our focus to methods recently used in light scattering studies. Table 1.4 shows several numerical methods widely used in recent light scattering simulation studies applied to aggregates.

Table 1.4. Methods widely used in recent light scattering studies of aggregates

1. Discrete dipole approximation (DDA) [16, 84, 85]
2. Finite difference time domain method (FDTD) [68, 69, 83]
3. T-matrix method for clusters of spheres (CTM) [42, 43]
4. Generalized multiparticle Mie-solution (GMM) [79, 80]
5. Effective medium theory combined with Mie theory (EMT+Mie) [12, 13]

In terms of the accuracy, we may divide the methods into two types: (1) rigorous and (2) approximate methods. Rigorous means that the solution for the numerical light scattering simulations are obtained rigorously, which produces final results with high accuracy. On the other hand, the approximate methods include approximations in the numerical light scattering simulations, which should be treated carefully to obtain results with the desired accuracy. We may group the CTM and GMM as rigorous methods, and the DDA, FDTD, and EMT as approximate ones.

1.3.1 The DDA and FDTD

The DDA and FDTD are methods based on volume integral formulations of Maxwell's equations. In the volume integral formulations, a particle is divided into an array of sub-volumes, then electromagnetic interactions between sub-volumes are treated in the simulations. Because of the treatment of the particle shape with sub-volumes, the DDA and FDTD can be used for various shapes of particles.

In the DDA, the particle is divided into an array of polarizable point dipoles. Electromagnetic interactions of the dipoles are calculated in the simulation with prescriptions for dipole polarizability (e.g., LDR; Lattice Dispersion Relation [16], SCLDR; Surface Corrected Lattice Dispersion Relation [14], GKLDR; LDR corrected in [29]).

The FDTD is based on replacement of temporal and spatial derivations of Maxwell's equations. The particle is divided into an array of small cubic cells. The FDTD uses the so-called ABL (absorbing boundary layer) to truncate the computational domain. The light interactions of the cells are calculated in the time domain. The vectors for electric and magnetic fields are treated at each discretized grid cell. Then, the time evolution of the electric and magnetic fields is simulated in the time domain until the state of convergence is obtained. Light scattering properties obtained in the time domain are converted into those in the frequency domain, that is, at different wavelengths (see [68, 69, 83] for further details of the FDTD).

In the DDA, errors are studied using the parameter $|m|(2\pi/\lambda)d$, where m is the complex refractive index, λ the wavelength of incident light, d the dipole separation. In the FDTD, errors are studied with the grid size described using λ (e.g., $\lambda/10$ and $\lambda/20$). The accuracy of the DDA and FDTD is often investigated by comparing them with those obtained using rigorous methods. For example, results for a compact sphere described with an array of sub-volumes are investigated by comparison with those obtained using Mie theory (e.g., [17, 18, 68]).

Draine and Flatau [19] recommended $|m|(2\pi/\lambda)d < 1.0$ for the efficiencies (e.g., Q_{ext}) and $|m|(2\pi/\lambda)d < 0.5$ for the phase function in order to suppress the errors less than a few percent for a compact sphere. For the FDTD, Sun and Fu [68] have shown that the grid cell size with $\lambda/20$ produces results with an error of less than 1% in Q_{ext} and Q_{sca} and of less than 5% in phase function. From these two papers and other related studies, it is shown that the size of sub-volumes should be less than a certain value to obtain results with the desired accuracy.

If the monomers are very small ($x_m \ll 1.0$), each spherical monomer can be represented with a single sub-volume (a dipole or a cubic cell) in the numerical simulations. On the other hand, when x_m becomes larger (e.g., $x_m \gtrsim 1.0$), each spherical monomer should be represented with an array of sub-volumes to obtain results with errors less than a few percent. However, the discretization of each monomer into an array of sub-volumes causes a limitation in the computable size parameter of the aggregate and requires longer computation time.

1.3.2 The CTM and GMM

In the CTM and GMM, incident electric field (E^{inc}) and scattered electric field (E^{sca}) are both expanded in vector spherical wave functions (VSWF; aka VSH; vector spherical harmonics) as in the following equations [43, 79]:

$$E_{inc} = \sum_{n=1}^{\infty} \sum_{m=-n}^{n} \sum_{p=1}^{2} l^i_{mnp} M_{mnp}, \tag{1.6}$$

$$E_{sca} = \sum_{n=1}^{\infty} \sum_{m=-n}^{n} \sum_{p=1}^{2} a^i_{mnp} N_{mnp}, \tag{1.7}$$

where m and n denote the order and degree, respectively. p denotes the mode of electromagnetic waves (i.e., TM and TE).

The relation between expansion coefficients of E^{inc} (l_{mnp}) and those of E^{sca} (a_{mnp}) is given by the T-matrix (translation matrix). Light scattering properties of the aggregate can be calculated from the T-matrix of the aggregate as described in [43]. In both the GMM and CTM, solutions for the analytical orientation averaging (AOA) are obtained [43, 79, 80].

The GMM and CTM are very similar in the formulation for a fixed direction of incident light. On the other hand, they are remarkably different in the formulations for the AOA [81]. The GMM deals directly with precise phase relations of both incident and scattered waves, while the CTM expresses both phase factors in terms of infinite series expansions [81]. The CTM is used for aggregates of spheres while the GMM can be applied to aggregates of non-spherical monomers on the condition

that the T-matrix of each monomer can be obtained. A comparative study between the GMM and CTM is given in [81].

The CTM and GMM are highly accurate because of the rigorous nature of the methods. However, the limitation exists in the largest value of monomer size parameter x_m, not because of the formulations, but because of the computational aspect. That is, the memory requirement increases very rapidly for the increase in x_m because higher orders of expansion coefficients in the VSWF become important in the light scattering properties for larger values of x_m. Also the memory requirement becomes larger for the increase in N.

1.3.3 The EMT

In the EMT, the light scattering properties of irregularly shaped particles having refractive index m are approximated by the light scattering properties of homogeneous particles (e.g., a compact sphere) having a so-called effective refractive index (m_{eff}) obtained from m and the porosity of the particles based on selected EMT formulations (e.g., the Maxwell–Garnett mixing rule and the Bruggeman rule) [13]. After calculating m_{eff}, Mie theory is used to obtain the approximate light scattering properties.

The simulations based on the EMT+Mie (i.e., the combination of the EMT and Mie) are much faster than other numerical methods for light scattering simulations. This, combined with its simplicity, has led to the EMT+Mie being widely used to study light scattering properties of irregularly shaped particles. However, the accuracy of the EMT+Mie needs to be considered.

Previous studies [73, 77, 78] have shown that Q_{sca} and Q_{abs} obtained from the EMT+Mie deviate greatly from those obtained using the DDA when the size parameter of voids or monomers is larger than around 2.0 (e.g., figs. 4–6 of [77]). The review of the EMT by Chylek et al. [13] concluded that errors in Q_{ext} and Q_{sca} are typically 10% and 15%, respectively, for x_m values up to 2.0.

The inaccuracy is caused by violation of the assumption in EMT formulations of negligible forward scattering by voids and inclusion particles [13]. The typical presence of forward-scattering peaks for larger voids and inclusions [8] invalidates the EMT method.

In addition to the inaccuracy of Q_{sca} and Q_{abs}, there are large inaccuracies in phase matrix elements obtained from the EMT+Mie. Therefore, this method should be used for small particles to obtain Q_{sca} and Q_{abs} and not phase matrix elements unless other types of EMT formulations are derived that take the forward scattering by voids and inclusions into account.

1.3.4 Future extensions of the numerical methods

As described above, various methods—both approximate and rigorous—can be applied to numerical light scattering simulations, depending on the desired accuracy, acceptable computation time, and available computer resources.

Related to the computation time, we need to treat the random orientation averaging of the aggregate. The CTM and GMM have the advantage that values of random orientation averaging can be obtained immediately after obtaining the

T-matrix of the aggregate. On the other hand, the DDA and FDTD are used for a fixed orientation of the aggregate. For random orientation averaging, simulations using the DDA and FDTD should be carried out repeatedly for the number of orientations required to reach convergence where further increase in the number of orientations does not influence the results.

Accurate numerical simulations of light scattering by realistic particles require the following improvements to the currently proposed methods:

1. Increasing the accuracy of approximate methods.
2. Speeding up simulations with approximate methods for a fixed orientation of the aggregate.
3. Speeding up procedures for the random orientation averaging.
4. Increasing the largest allowable monomer size in both approximate and rigorous methods.
5. Increasing the treatable number of monomers of the aggregate in both approximate and rigorous methods.

1.4 Improved numerical simulations

1.4.1 Grouping and adding method (GAM)

One of the limitations in the numerical light scattering simulations of aggregates is the largest number of monomers N, which can be treated. N is limited because of memory and computation time. The computation time can be reduced with the advent of workstations having higher performance (e.g., multiple core system). On the other hand, the memory limitation is sometimes difficult to solve even with the evolution of computer hardware, because the higher N is, the higher the orders of scattering functions that should be treated in the numerical light scattering simulations.

In this circumstance, we have proposed a method to reduce the maximum memory required in numerical light scattering simulations. We call the method GAM (Grouping and Adding Method) [58]. Following is the procedure of the GAM:

1. Divide the target aggregate into N_{DIV} groups of sub-aggregates
2. Calculate efficiencies (Q_{sca} and Q_{abs}) of each sub-aggregate
3. Sum the efficiencies of all the sub-aggregates

The GAM has a feature to reduce required memory by treating one group of sub-aggregates at one time. Because of this division, we can carry out simulations for larger aggregates, which originally would have required memory larger than our computer resources.

The GAM is combined with a fixed orientation version of numerical methods for light scattering simulations. We used a fixed orientation version of the CTM [43].

The summation of light scattering properties of all the groups is possible. This is because the total efficiencies of an aggregate can be obtained based on superposition

principle of those from all the monomers as shown in the following equations [42]:

$$C_{ext}^{agg} = \sum_{i=1}^{N} {C_{ext}}^{i}, \tag{1.8}$$

$$C_{abs}^{agg} = \sum_{i=1}^{N} {C_{abs}}^{i}, \tag{1.9}$$

where C denotes the cross-section. agg denotes cross-sections of the aggregate. i represents a contribution from ith monomer.

The GAM takes advantage of the feature that monomers with larger x_m start to have strong forward scattering (see Fig. 1.3). With the strong forward scattering by one monomer (hereafter called the ith monomer), light incident on the ith monomer will be scattered mostly into the forward direction. Then, monomers located at a distance in the side direction from the ith monomer rarely contribute to the multiple interactions of the light, which is first scattered by the ith monomer.

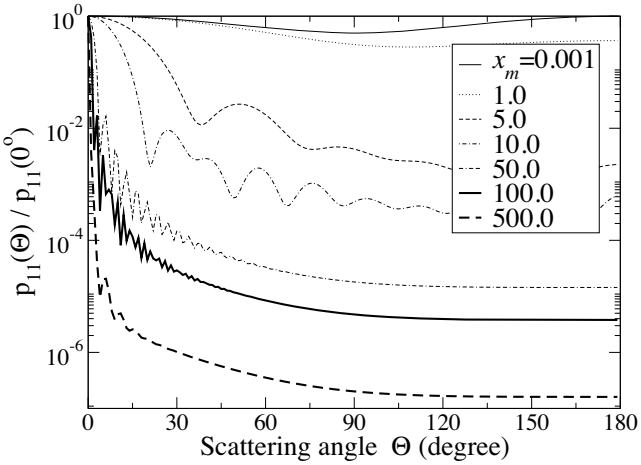


Fig. 1.3. Scattering phase function $p_{11}(\Theta)$ of a single monomer as a function of the size parameter of the sphere, x_m, where $m = 1.5+0.1i$. Each scattering function is normalized to its value at $\Theta = 0^\circ$ (after Y. Okada et al., *JQSRT*, **108**, 65–80, 2007). p_{11} is obtained from a_1 shown in section 1.2.1

Fig. 1.4 shows schematic figures of the grouping of an aggregate. The division of the aggregate is carried out based on a plane perpendicular to the direction of incident light. With this way of division, we can take advantage of the forward scattering feature occurring for larger monomer size. For the grouping, we use the K-means method [45].

When we calculate light scattering properties of one group, we treat monomers of the group, together with monomers of surrounding regions. We call the surrounding region the buffer region (see Fig. 1.4(b)). Table 1.5 shows the parameters of the GAM.

The light scattering properties of the buffer region (N_{gr}^{out} monomers) are not obtained accurately because of the absence of monomers outside the buffer region ($N_{org} - N_{gr}^{out} - N_{gr}^{in}$). Therefore multiple scattering between monomers in the buffer regions and those in outer surrounding regions becomes insufficient compared with when the aggregate is not divided into groups.

For each group, we carry out one simulation using N_{gr}^{in} monomers of the group and N_{gr}^{out} monomers of the buffer region. Then we store the light scattering properties only of N_{gr}^{in} monomers for future summations of all the groups. We discard results of N_{gr}^{out} monomers of the buffer region because of the inaccuracy. However, the monomers selected as the buffer region (N_{gr}^{out}) in the simulation of one group will be treated as monomers of the group (N_{gr}^{in}) in the simulation of another group. Then in the summation of the results of all the groups, we can obtain results contributed from all the monomers.

Table 1.5. Parameters of the GAM

N_{org}	: Number of all the monomers of the original aggregate
N_{DIV}	: Number of groups
N_{gr}^{in}	: Number of monomers in the selected group
N_{gr}^{out}	: Number of monomers in the buffer region

The GAM will cause errors because the absence of monomers outside the buffer region will exclude the influence of ($N_{org} - N_{gr}^{in} - N_{gr}^{out}$) monomers on multiple light scattering of N_{gr}^{in} monomers. However, we expect that those errors will become smaller for larger x_m because of forward scattering characteristics. We have investigated the errors by using aggregates treatable with memory size of our computers (2GB) for $N = 512$ and x_m up to around 4.5.

The errors are estimated by comparing between efficiencies of the aggregate (1) with the GAM and (2) without the GAM:

$$Error(\%) = \frac{|Q^{GAM} - Q^{org}|}{Q^{org}} \times 100, \tag{1.10}$$

where GAM is the results with the GAM. org denotes results without the GAM.

The size of the buffer region influences the accuracy of light scattering properties of N_{gr}^{in} monomers. Errors caused by the GAM are investigated for different values of the ratio N_{gr}^{out}-to-N_{org}. It is found that an N_{gr}^{out}-to-N_{org} of 1/8 will have errors of up to 15% for x_m of up to 4.5, $N_{org} = 512$, $N_{DIV} = 60$ and for various sets of (n, k) (see Fig. 1.5). For absorbing monomers (i.e., higher k values), the ratio N_{gr}^{out}-to-N_{org} can be taken as 1/16 for errors of up to 15%.

For errors for different values of N_{DIV} and for other details of the GAM, refer to our paper [58]. Sample code using the GAM is available at [88].

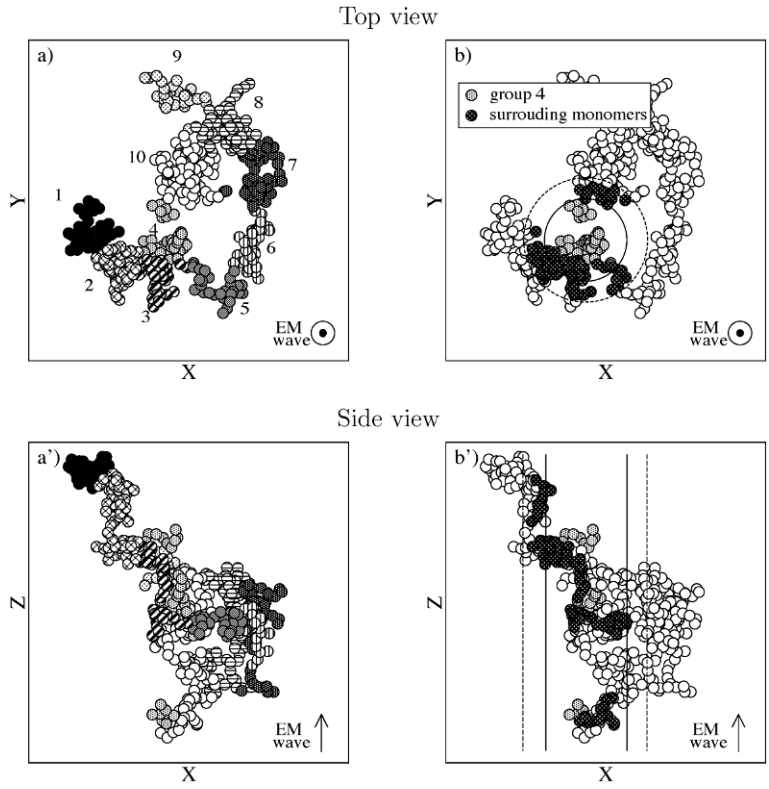


Fig. 1.4. Schematics of grouping of the aggregate (a and a′) and of the buffer region (b and b′) for EMs propagating in the Z direction. The original aggregate (BCCA512) is grouped with $N_{DIV} = 10$. In panels b and b′, gray and black shaded circles are the N_{gr}^{in} monomers in the group and N_{gr}^{out} monomers in the buffer region, respectively, and the large solid circle and dashed one denote the maximum radius of group 4 and buffer region for the group, respectively. The direction of the incident light is shown at the bottom right of each panel (after Y. Okada et al., *JQSRT*, **108**, 65–80, 2007).

1.4.2 Numerical orientation averaging using a quasi-Monte-Carlo method (QMC)

Light scattering properties of random orientation averaging of an aggregate are obtained by integrating results over a large number of orientations analytically or numerically (see section 1.2.3). Numerical orientation integration of results using fixed orientation code is traditionally performed by using multiple sets of orientations changed with some certain steps in each of three Euler angles. Hereafter we call the division of orientations with certain steps the LAT (lattice grid division).

Computational efficiency of the numerical orientation integration depends on how multiple orientations are selected (i.e., their number and their distribution in angular space). The LAT is easy to treat; however, it is not necessarily the best way to efficiently carry out the integration.

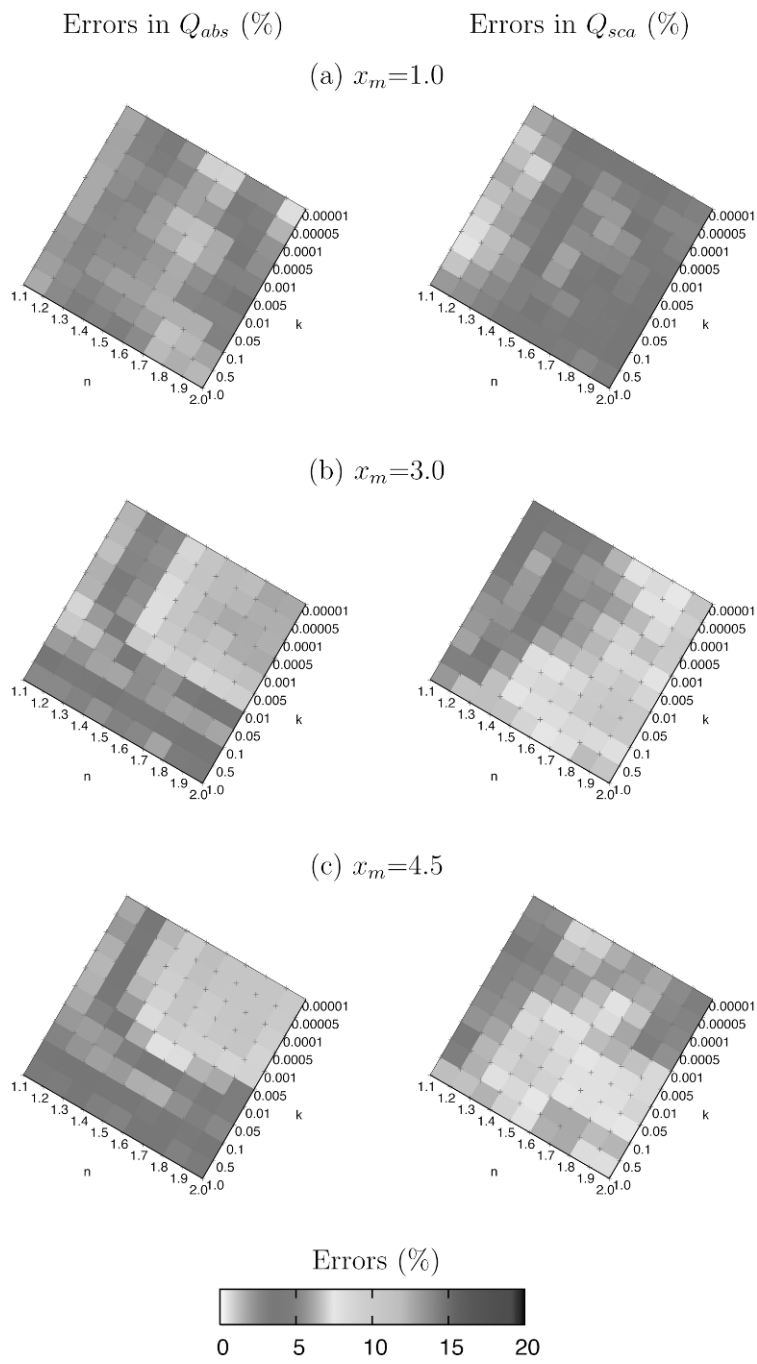


Fig. 1.5. Maps of errors in Q_{abs} and Q_{sca} for n vs k for different x_m. The buffer region has an N_{gr}^{out}-to-N_{mono} ratio of 1/8, and $N_{mono} = 512$ and $N_{DIV} = 60$ (after Y. Okada et al., *JQSRT*, **108**, 65–80, 2007).

Efficient methods for integrations over the 1-D, 2-D, and 3-D have been studied, which resulted in a so-called QMC (Quasi Monte–Carlo) method for an efficient integration. We have proposed to apply the QMC for random orientation averaging of the light scattering properties in [59].

The QMC is similar to a Monte–Carlo method using random numbers, which is conventionally called the CMC (Crude Monte–Carlo) method [57]. By using the CMC or QMC, the orientation integration is carried out as the following equation:

$$\frac{1}{A}\int_A f(s)\,ds = \frac{1}{N_{int}}\sum_{i=0}^{N_{int}-1} f(x_i), \tag{1.11}$$

where A denotes the area of the considered region, ds the area element in A, and x_i the position of the ith point, $f(x_i)$ is the value of the function at this point, and N_{int} the number of points used for the integration.

It is known that the CMC converges with $O(\sqrt{\log\log N}/\sqrt{N})$ [57]. Hence the accuracy of the CMC increases only slowly as $O(1/\sqrt{N})$.

The QMC is developed to increase the convergence efficiency compared with the CMC. In the QMC, instead of the random numbers used in the CMC, a number sequence called the LDS (Low-Discrepancy Sequence) is used.

The LDS comprises a uniform number sequence between 0 and 1. Among various LDSs, we used the so-called HS (Halton sequence) for two-dimensional variables. The HS is generated by a one-dimensional van der Corput sequence. The van der Corput sequence $g_b(k)$ can be recursively obtained as [57]:

$$g_b(0) = \frac{\sigma(0)}{b-1}, \tag{1.12}$$

$$g_b(bn+r) = \frac{1}{b}(g_b(n)+\sigma(r)), \tag{1.13}$$

where b is the base, and σ a permutation of natural numbers ranging from 0 to $b-1$ (i.e., $\sigma = \{0, 1, \ldots, b-1\}$).

For example, a van der Corput sequence with base of 3 (i.e., $g_3(k)$ with $b=3$) can be calculated as follows. $g_3(0) = \sigma(0)/(3-1) = 0/(3-1) = 0$ according to Eq. (1.12). $g_3(1)$ is considered as $g_3(3*0+1)$ with $n=0$ and $r=1$ in Eq. (1.13). Then, $g_3(3*0+1) = \frac{1}{3}(g_3(0)+\sigma(1)) = \frac{1}{3}(0+1) = \frac{1}{3}$. Recursively we can obtain $g_3(k) = \{\frac{1}{3}, \frac{2}{3}, \frac{1}{9}, \frac{4}{9}, \ldots\}$ for $(k = 1, 2, 3, 4, \ldots)$.

The HS for 2-D is produced by a combination of two van der Corput sequences with two different bases. For example, we can use 2 and 3 as bases for the two orientation angles (α and β) of the Euler angles. The number sequence ranging between 0 and 1 produced from the HS is converted into the full range of the angle (e.g., 0 to 2π) used for the orientation rotations.

Fig. 1.6 shows the difference in orientation distributions using above methods. In Fig. 1.6 we compare:

1. CMC using random number generator (*rand*) commonly implemented in Fortran compilers,
2. CMC using MT (Mersenne Twister) [44],
3. QMC using HS with bases of 2 and 3.

The one prominent difference between *rand* and the MT is the length of periodicity of random numbers. The periodicity of the MT is much longer than that of *rand*, making the CMC with the MT more secure for use in integrations of higher dimensional space. However, the above two methods using the CMC show point distributions whose uniformity is not as good as those of the QMC-based method (see Fig. 1.6). This difference in the uniformity of orientation distribution causes a difference in convergence efficiencies among these three methods.

The efficiency of convergence of the QMC is mathematically known as $O[(\log N)^s/N]$ where s is dimension (e.g., $s = 2$ for two angles) [57]. Therefore the convergence efficiency of the QMC ($\propto O(1/N)$) is better than that of CMC ($\propto O(1/\sqrt{N})$).

Fig. 1.7 shows the comparison of convergence of the light scattering properties as a function of the number of orientations between three methods for orientation selections for the NOA (i.e., LAT, CMC with MT, and QMC with HS). We used a fixed orientation version of the CTM to obtain the light scattering properties of each orientation of the aggregate. Only phase functions at several scattering angles of BPCA128 (i.e., BPCA composed of 128 monomers) are shown (see [59] for comparisons of other light scattering properties, and for other shapes of aggregates such as bisphere, BCCA4, and BCCA128).

The bold gray lines in Fig. 1.7 are results obtained using the AOA version of the CTM (see Table 1.4). Therefore, the approach of results to the bold lines for an increased number of orientations denotes the convergence of the NOA.

The convergence efficiencies are better in the QMC followed by the CMC. The convergence efficiency of the LAT is the worst among the three methods for the phase function shown in Fig. 1.7. The convergence efficiencies depend on types of light scattering properties (e.g., Q_{abs}, polarization at some scattering angles, and g); however, the general tendency has the same trends as shown in Fig. 1.7.

From these studies, it is shown that the QMC method can be used for efficient numerical orientation averaging of the light scattering properties of aggregates. Sample Fortran code to generate the HS is available at [89].

1.4.3 Extended calculation of light scattering properties with numerical orientation averaging

x_m and N of an aggregate treatable in the light scattering simulations are limited mainly because of the computer memory. For example, when we use the AOA version of the CTM on computers with 2 GB memory, the maximum value of N becomes around 200 for $x_m \approx 1.7$. Because of these limitations in the CTM and those in other methods for numerical light scattering simulations (e.g., the DDA, the GMM), studies of the light scattering properties of fractal aggregates have been confined to a certain range of parameter space. Fig. 1.8 shows the parameter space (N and x_m) considered in previous studies of the light scattering properties of aggregates.

In Fig. 1.8, it is shown that values of N with $x_m \approx 1.7$ are confined to be less than around 200 in most studies. This limitation in parameter space clearly limits our understanding of the light scattering properties of aggregates. Also it is not clear whether or not the light scattering properties of the aggregates with

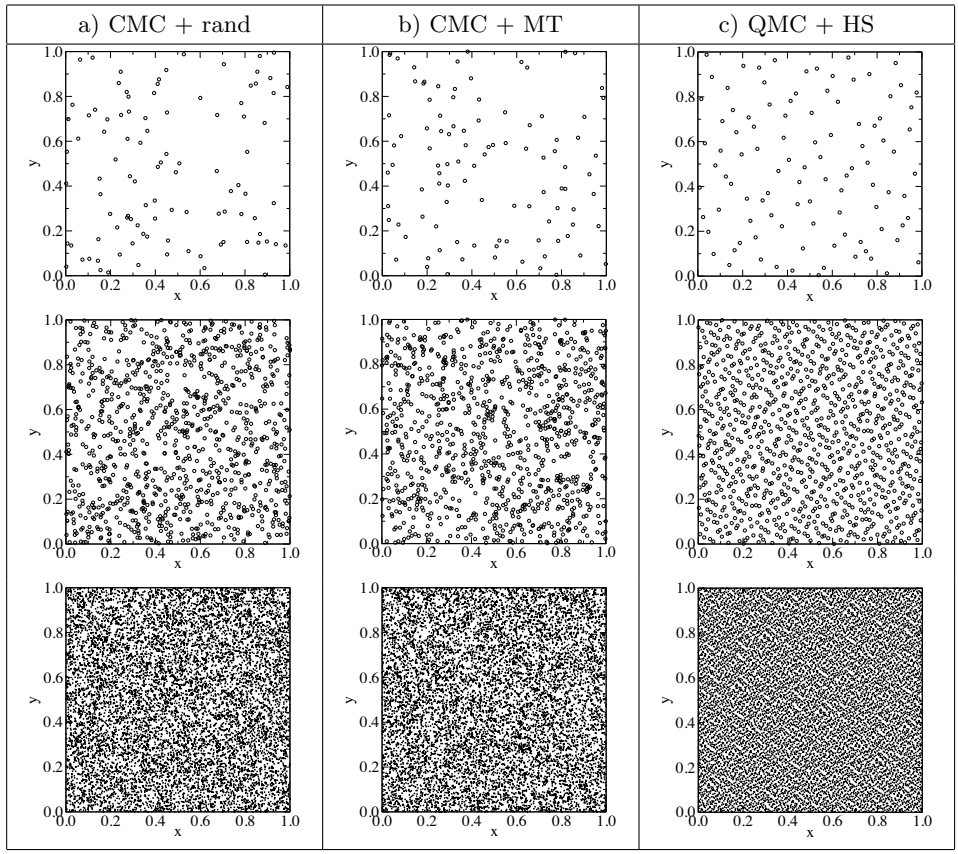


Fig. 1.6. Comparison of distribution of two variables (x and y). (a) CMC+rand: crude Monte–Carlo method with random number generator (*rand*()) implemented in Fortran compiler. (b) CMC+MT: crude Monte–Carlo method with Mersenne Twister. (c) QMC+HS: quasi Monte–Carlo method with Halton sequence. From top to bottom show results for 100, 1000 and 10 000 points (after Y. Okada, *JQSRT*, **109**, 1719–1742, 2008).

$N \lesssim 200$ are representative of those of aggregates realistically existing in nature (e.g., cometary dust, microbiocells, soot aerosols). It is important to investigate the validity of these light scattering properties by comparing them with results of aggregates having increased N.

In order to increase N and x_m treatable in the numerical light scattering simulations, we have proposed to use a fixed orientation version of the CTM combined with the NOA instead of the AOA in the random orientation averaging.

Public Fortran codes of the fixed orientation version and AOA version of the CTM are available on the Web [87]. We used two Fortran codes (i.e., scsmfo1b.for and scsmtm1.for) for the fixed orientation version and the AOA version of the CTM, respectively.

The two versions of the CTM are different in memory usage because the former treats only the fixed orientation. Also this version uses optimization for memory

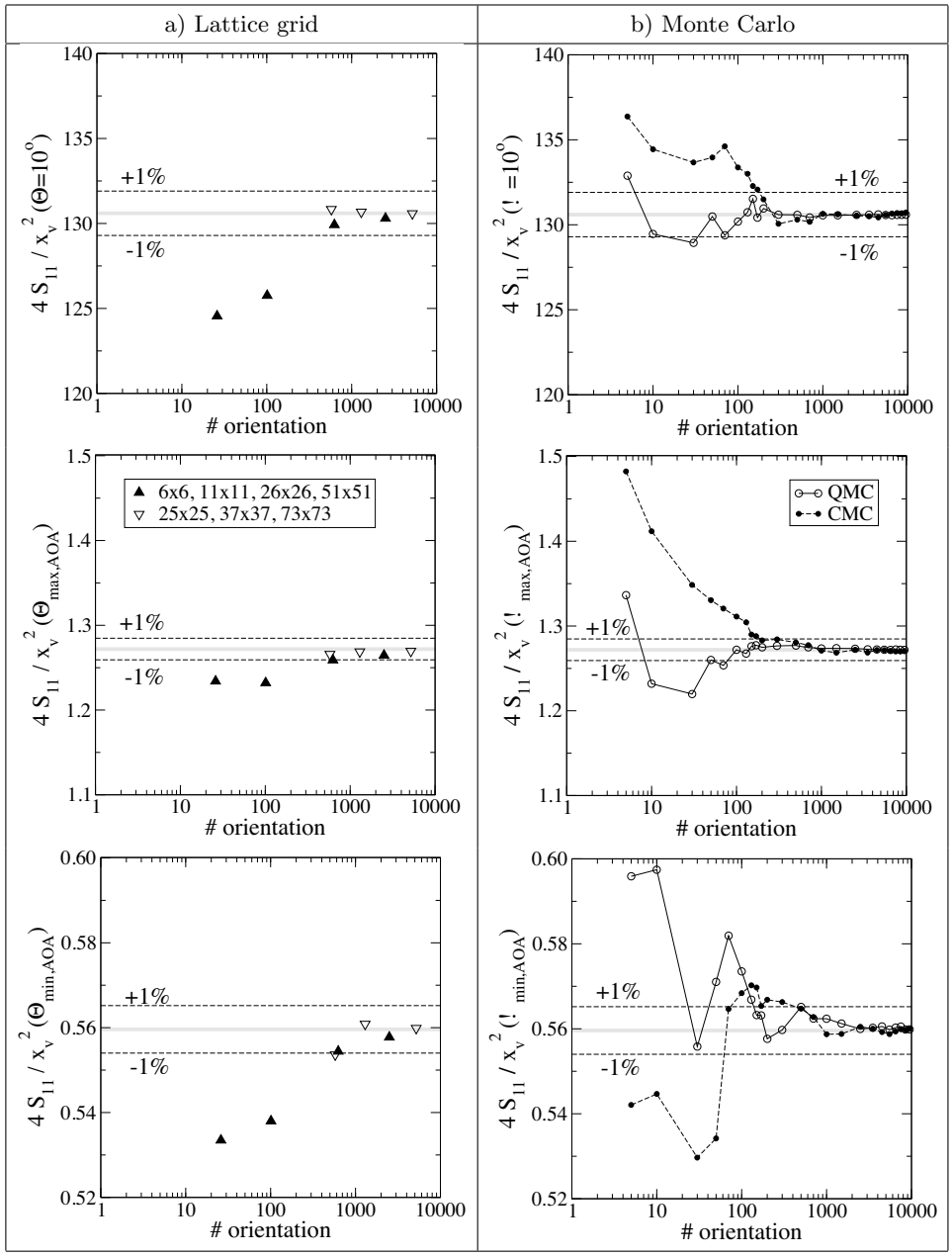


Fig. 1.7. Comparison of intensities at scattering angle of 10 degrees obtained using three different methods for the NOA. S_{11} denotes phase function (i.e., a_1 in section 1.2.1). x_v denotes the size parameter of a sphere with equivolume to the aggregate. X axis denotes number of orientations (i.e., a combination of zenith and azimuth angles) used for the NOA. Gray lines are values obtained from the AOA. Two dashed lines are error bounds estimated from the value of the AOA. (After Y. Okada, *JQSRT*, **109**, 1719–1742, 2008)

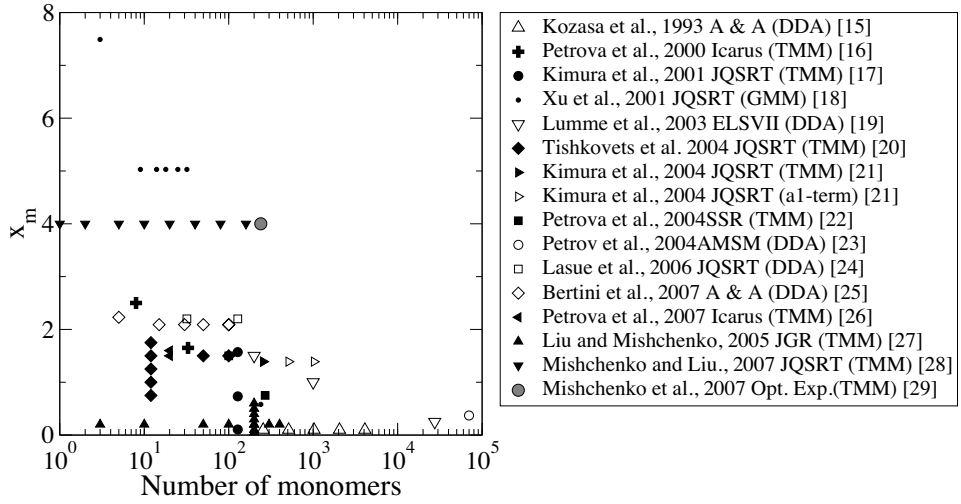


Fig. 1.8. Parameter space in past studies of the light scattering properties of fractal aggregates. Note that the reference numbers on the right-hand side are those in the original paper (after Y. Okada et al., *JQSRT*, **109**, 2613–2627, 2008).

usage by compaction into the amplitude matrix devised for the fixed orientation version (see a document found at [87]). On the other hand, the AOA version requires memory for random orientation averaging, which is much larger than that required for the fixed orientation version. This difference in the memory usage between the two versions of the CTM will cause a difference of the upper limitations in N and x_m.

Fig. 1.9 shows the improvements of the parameter space by using the fixed orientation version of the CTM (scsmfo1b.for) from those of the AOA version of the CTM (scsmtm1.for). It is clear that the fixed orientation version of the CTM increases N and x_m greatly compared to the AOA version of the CTM. When we use the fixed orientation version, we need to carry out the random orientation averaging numerically using the NOA. As introduced in section 1.4.2, we recommend the QMC for an efficient random orientation averaging with the NOA.

Some results of numerical light scattering simulations based on this technique for a medium comprising spherical particles are given in [61].

A software package named NAOPC (numerical averaging of optical properties of clusters) using the techniques in this section and section 4.2 are available at [90].

1.4.4 Scattering and absorption of BCCA composed of tens to thousands of monomers

We introduce light scattering properties of BCCA composed of tens to thousands of monomers using our proposed methods. The fixed orientation version of the CTM can be used for $N \lesssim 512$ when we consider monomers of $x_m = 4.5$ on computers with 2 GB RAM. As written in section 1.4.1, the maximum N can be increased by the GAM. Errors caused by the GAM are less for absorbing materials than those

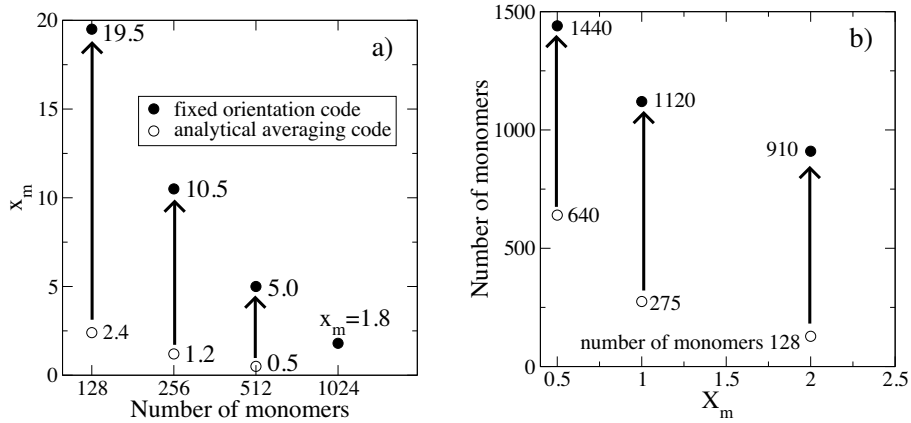


Fig. 1.9. Improvements in parameter space by using fixed orientation version of the CTM instead of the AOA version of the CTM (After Y. Okada et al., *JQSRT*, **109**, 2613–2627, 2008).

of low to intermediately absorbing materials. Therefore, for absorbing materials, we can use the GAM with the ratio N_{gr}^{out}-to-N_{org} of 1/16, while for other refractive indices we used the ratio of 1/8 to obtain results with errors of less than 15%.

We consider three materials: (1) silicate, (2) pyroxene, and (3) amorphous carbon (see Table 1.6 for their refractive indices). The pyroxene is actually one type of silicate. However, we call them silicate and pyroxene to denote low-absorbing and intermediately absorbing materials, respectively.

Table 1.6. Refractive index of materials

Silicate:	$m = 1.48 + 0.000028i$ [54]
Pyroxene:	$m = 1.635 + 0.0042i$ [15]
Amorphous carbon:	$m = 2.03 + 0.77i$ [66]

Fig. 1.10 shows Q_{abs} and Q_{sca} obtained from our simulations. Results with $N \leq 512$ are obtained using the CTM without the GAM; those with $N \geq 1024$ are obtained using the CTM combined with the GAM.

Note that the simulation results are obtained for a single fixed orientation of the aggregate, not for random orientation averaging.

Results clearly show a linear relationship between N and efficiency values (i.e., Q_{abs} and Q_{sca}). In results with N of 128 and 256, there are some deviations from the linear relationship. However, results are obtained for a single orientation, not random orientation averaging. We consider that these deviations may be caused by specific alignment of monomers for a single orientation, and may disappear for random orientation averaging.

We obtained the regression of the relations between N and efficiency values (Q_{abs} and Q_{sca}). The obtained regression equations are shown in Table 1.7. The correlation coefficient of each linear regression line is shown in brackets in Table 1.7.

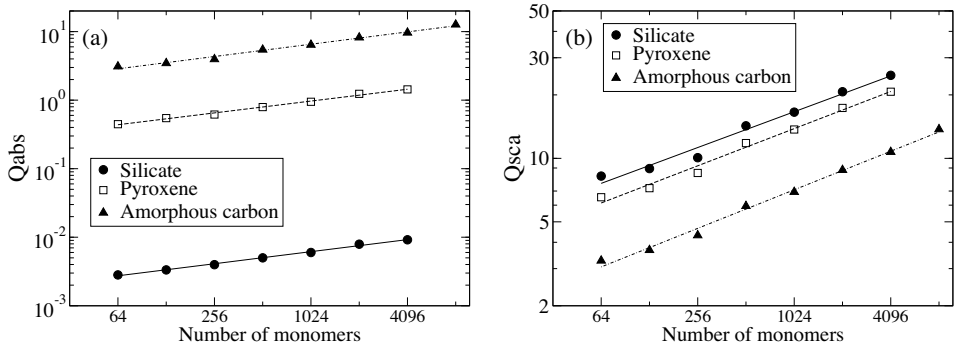


Fig. 1.10. Log–log plots of Q_{abs} and Q_{sca} of aggregates composed of up to 8192 monomers, where $x_m = 4.5$ for silicate, pyroxene, and amorphous carbon. The lines are the regression line for each material. The aggregates with $N \leq 512$ are calculated without the GAM while those with $N \geq 1024$ are calculated with the GAM. All the results are obtained for one orientation of the aggregate. Table 1.7 lists the equations of the regression lines (after Y. Okada et al., *JQSRT*, **108**, 65–80, 2007).

All the correlation coefficients show very high values (i.e., $\geq$0.990) denoting strong correlation between N and efficiency values for all the materials.

It is also noted that exponential slopes of regression lines of different materials are very similar (i.e., 0.293 ± 0.008). Therefore, we consider that the relationship between N and efficiency values does not depend on the material composition, rather it depends on other physical properties of the BCCA; one possibility is the shape of the aggregate.

Other details of the simulations are discussed in [58].

Table 1.7. Equations of regression lines in Fig. 1.10 for silicate (Si), pyroxene (Py) and amorphous carbon (Ac). The numbers in brackets represent the correlation coefficient of the regression, and y represents the efficiencies (i.e., Q_{abs} and Q_{sca}), respectively (after Y. Okada et al., *JQSRT*, **108**, 65–80, 2007).

	Q_{abs}		Q_{sca}	
Si	$y = 0.000811 \times N^{0.292}$	(0.998)	$y = 2.35 \times N^{0.282}$	(0.990)
Py	$y = 0.133 \times N^{0.287}$	(0.997)	$y = 1.82 \times N^{0.293}$	(0.994)
Ac	$y = 0.831 \times N^{0.298}$	(0.995)	$y = 0.864 \times N^{0.304}$	(0.996)

1.4.5 Intensity and polarization of light scattered by silicate aggregates

The intensity and polarization of light scattered by fractal aggregates are important light scattering properties utilized to interpret observational data of cometary dust. Numerical light scattering simulations have been utilized to study the light scattering properties of cometary dust using fractal aggregates as their shape models

(e.g., [1,2,34,35,37,39,62–64,70,71]). However, as shown in section 1.4.3, aggregates studied before have been limited by $N \lesssim 200$ with $x_m \approx 1.7$. It is important to study light scattering properties of the aggregates with increased N because realistic dust particles are considered to have a much larger number of sub-components (e.g., one million) as deduced from the image of IDP (Interplanetary dust particles) (see fig. 3 of [30], for example).

In this part, we show sample results of the light scattering properties of the BPCA and BCCA with increased N obtained from our extended numerical light scattering simulations [60]. We consider the wavelength λ of 0.6 μm, and refractive index of silicate $m = 1.68 + 0.03i$ [38]. We treat the BPCA and BCCA composed of 128, 256, 512, and 1024 monomers. We consider different values of monomer radius r_m. Results with $N \geq 512$ can be obtained by using a fixed orientation version of the CTM on our computer resources (2 GB RAM). We investigate results of random orientation averaging obtained using the NOA procedure with the fixed orientation version of the CTM.

Figs 1.11 and 1.12 show the intensity ($a_1(\Theta)$ in Eq. (1.1)) and polarization ($P(\Theta)$ in Eq. (1.3)) of the BPCA and BCCA with N ranging from 128 to 1024. Curves with different styles correspond to results with different N.

Intensity

The results of intensity (i.e., left panels of Figs 1.11 and 1.12) show the following two impacts of N:

- Forward scattering amplitudes at scattering angle of 0° for the BPCA and BCCA increase with N. The BCCA shows a larger increase of the amplitudes with N compared with that of the BPCA.
- A bump around the scattering angles of 10–20° becomes noticeable for the BPCA with increased N (i.e., 256–1024). The scattering angle of the bump shifts from 20° to 10° as N increases. Also the bump becomes sharper for the BPCA with increased N. In the case of the BCCA, the bump is not seen even with $N = 1024$.

We consider that the increase of forward scattering amplitudes of the BPCA and BCCA is caused by increased multiple light scattering between monomers. Because of the difference in the structures of the BPCA and BCCA, the influence of N on the intensity is different.

There are several mechanisms of multiple light scattering between monomers, such as CB (coherent backscattering) effect [53], SH (shadow hiding) effect [67], near-field effect [65]. These different mechanisms work differently for aggregates having different structures.

As introduced in section 1.2.2, the BPCA has a porosity of around 85% while the BCCA has a porosity greater than 99%. That is, the BPCA is more compact compared with the BCCA.

The second feature of the intensity (i.e., the bump around scattering angles of 10–20°) becomes clearly noticeable for the BPCA with larger N (e.g., 1024). Therefore, we shall consider that the bump is caused by a mechanism which works effectively for compact aggregates (BPCA) with higher orders of multiple scattering, and which does not work for sparse ensembles (BCCA).

One possible mechanism causing above two intensity features is the near-field effect. According to Petrova et al. [65], the near-field effect works more for compact structures while the CB effect works effectively for sparse ensembles. Future studies of different mechanisms of multiple light scattering inside an aggregate are interesting to further understand results of numerical light scattering simulations of aggregates.

Polarization

Middle and right panels of Figs 1.11 and 1.12 show degree of polarization by the BPCA and BCCA. The right panels represent polarization values at scattering angles of 150–180° in order to concentrate on the negative branch of the polarization.

Hereafter we use symbols listed in Table 1.8 for discussion of the polarization characteristics. Influence of N on the polarization is as follows:

1. Influence of N on P_{max}:
   (a) P_{max} at scattering angles of around 90° decreases as N increases.
   (b) The increase of N greatly reduces P_{max} for the BPCA while it reduces P_{max} little for the BCCA.
2. Influence of N on the NPB (negative polarization branch):
   (a) Θ_{min} shifts to larger scattering angles as N increases. This influence is commonly seen for the BPCA with r_m of 0.11, 0.13, and 0.15 μm.
   (b) P_{min} changes for different values of N. This influence is clear for the BPCA and not clear for the BCCA.

The decrease of P_{max} for increased N (items 1(a) and 1(b) of the above list) is explained by the depolarization effect due to the increased orders of multiple light scattering between monomers (see section 13.4 of [50]). Since the BPCA is more compact than the BCCA, we consider the influence of the depolarization to be larger for the BPCA than the BCCA.

As regards item 2(a) of the above list, Θ_{min} shifts from around 165° to around 175°. This shift is less pronounced for the BPCAs with $r_m = 0.11$ and 0.13 μm, but quite clear for the BPCA with $r_m = 0.15\,\mu$m.

Volten et al. [72] have performed laboratory measurements of polarization obtained for several fluffy aggregate samples. The polarization data are obtained for the scattering angles up to around 175°.

The polarization curves in [72] have a feature that the value of P decreases for larger scattering angles. The minimum value of the polarization is located at the largest scattering angle attainable in their laboratory measurements (i.e., 175°). Θ_{min} was not clear in the measurements because the largest scattering angle was limited. However, we can expect from their measurements that Θ_{min} will be located at scattering angles of 175–180°. This similar trend of Θ_{min} between laboratory measurements and numerical light scattering simulations is interesting for future study of the NPB.

Now we focus on item 2(b) of the above list. With the increase of N, P_{min} changes differently for BPCAs having different values of r_m:

- P_{min} monotonously increases for the BPCA having r_m of 0.11 μm,
- P_{min} temporarily increases, then decreases for the BPCA having r_m of 0.13 μm
- P_{min} monotonously decreases for the BPCA having r_m of 0.15 μm

The NPB is generated by some mechanisms of multiple light scattering between monomers (e.g., CB, SH, and near-field effect). Properties of multiply scattered light depend on how each monomer scatters the light impinging on it. We consider that monomers of BPCAs having different r_m somewhat differently scatter the impinging light causing the different behavior of the NPB with increase of N.

Compared to the BPCA, the BCCA does not show a clear influence of N on the NPB. The right-hand panels of Fig. 1.12 show a small variation in the NPB for different values of N. However, the variation is not monotonous for the increase of N. We do not currently know whether this variation is actually caused by the different values of N or by other variation occurring during the generation of the BCCA (e.g., a small-scale variation between BCCAs having the same values of N).

We need further investigation on the influence of N and other physical properties on the intensity and polarization of the aggregates to better understand how the several mechanisms of multiple light scattering inside an aggregate work.

Table 1.8. Symbols of the polarization characteristics

The maximum value of the polarization:	P_{max}
The scattering angle of the P_{max}:	Θ_{max}
The minimum value of the polarization seen at the NPB:	P_{min}
The scattering angle of the P_{min}:	Θ_{min}

The NPB denotes a negative polarization branch at backscattering angles.

1.5 Summary

In this chapter, we have introduced methods for numerical light scattering simulations applicable to aggregates whose size is small to moderate compared to the wavelength of the incident light.

Methods developed around the 1970s as shown in section 1.3 were applicable to aggregates with a small number of monomers (e.g., five monomers) and to those with small monomers in the Rayleigh region. Those methods were later improved to be applicable to aggregates with a larger number of monomers and with increased monomer sizes. Also in section 1.3, the methods widely used in recent numerical light scattering studies of aggregates were briefly described. These methods are applied to aggregates composed of hundreds of spheres when the monomer size parameter is less than around 2.0. Further improvements of these methods are required for the full application of light scattering properties obtained from the numerical simulations for the interpretation of observational data.

In section 1.4, we presented techniques for improving the numerical light scattering simulations published in recent years. They include:

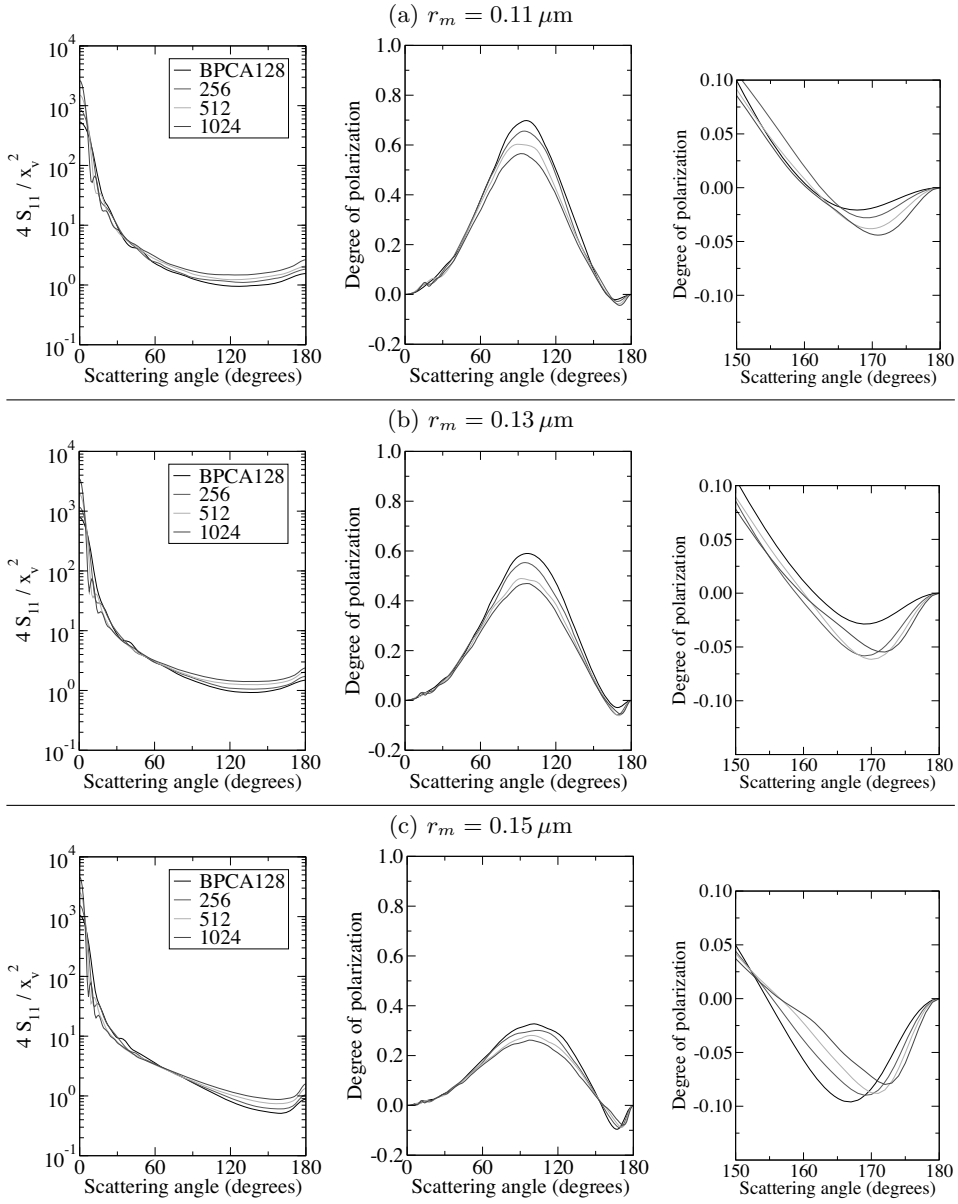


Fig. 1.11. Intensity and polarization of BPCA for various numbers of monomers. Left and middle panels: intensity and polarization at all the scattering angles. Right panels: Polarization in negative polarization branch (After Y. Okada et al., *JQSRT*, **109**, 2613–2627, 2008).

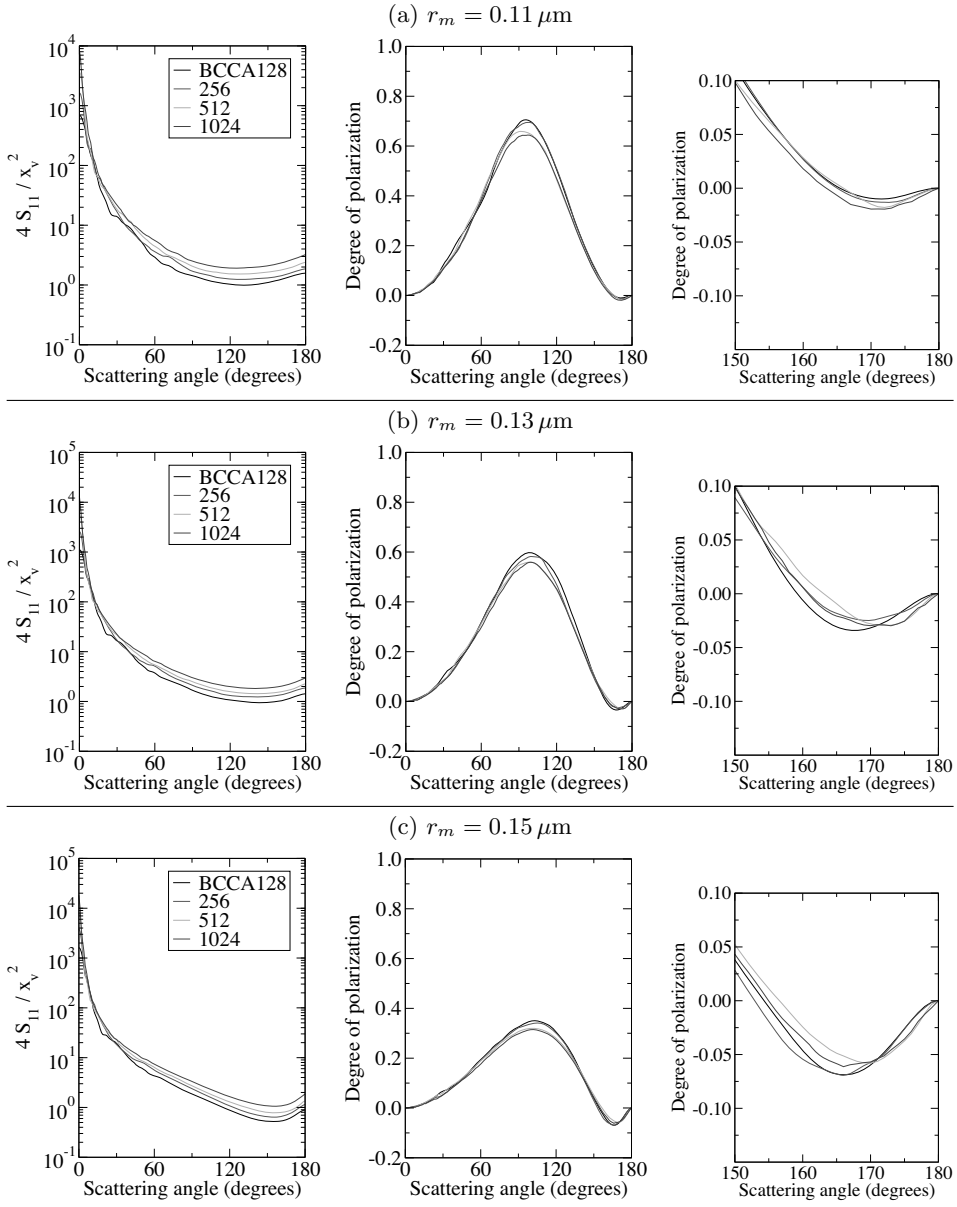


Fig. 1.12. Same as Fig. 1.11 but for BCCA.

- the grouping and adding method (to increase number of monomers),
- the quasi Monte–Carlo method (to efficiently carry out random orientation averaging),
- Numerical orientation averaging (to extend treatable parameter space, i.e., the size and number of monomers).

Improvements in numerical light scattering simulations enable us to improve understanding of the light scattering properties of aggregates, and also to interpret the observational data of various objects having an aggregate structure (e.g., dust in space, microbiocells, and atmospheric particles).

Acknowledgments

Studies introduced in this article used cluster T-matrix code developed by D. Mackowski, K. Fuller, and M. Mishchenko. Two types of code (for the AOA and for a fixed orientation) are obtained from http://www.giss.nasa.gov/˜crmim/. We acknowledge T. Mukai and I. Mann for their comments on an earlier version of this chapter. We appreciate reviewers for their suggestions to improve this chapter. Credit also goes to A. A. Kokhanovsky for his invitation to us to prepare this chapter and stimulating discussions. Some of the data analyses were performed on the general common use computer system at the Astronomical Data Analysis Center (ADAC) of the National Astronomical Observatory of Japan. The author was supported by the Interdisciplinary Graduate School of Science and Engineering, Kinki University, Osaka, Japan, during preparation of the manuscript. The present work was mainly carried out with the support of The 21st Century COE Program of Origin and Evolution of Planetary Systems (2003 Oct.–2008 March) of the Japanese Ministry of Education, Culture, Sports and Technology (MEXT).

Abbreviations

AOA : analytical orientation averaging
BCCA : ballistic cluster–cluster aggregate
BPCA : ballistic particle–cluster aggregate
CB : coherent backscattering
CMC : crude Monte–Carlo
CTM : T-matrix method for clusters of spheres
DDA : discrete dipole approximation
EMT : effective medium theory
FDTD : finite difference time domain method
GAM : grouping and adding method
GMM : generalized multiparticle Mie-solutions
HS : Halton sequence
LAT : lattice grid division
LDS : low-discrepancy sequence
MT : Mersenne Twister
NOA : numerical orientation averaging

NPB : negative polarization branch
QMC : quasi-Monte–Carlo
SH : shadow hiding
VSWF : vector spherical wave function
VSH : vector spherical harmonics

References

1. Andersen, A.C., H. Mutschke, T. Posch, M. Min, and A. Tamanai, 2006: Infrared extinction by homogeneous particle aggregates of SiC, FeO and SiO2: Comparison of different theoretical approaches, *J. Quant. Spectrosc. Rad. Transfer*, **100**, 4–15.
2. Bertini, I., N. Thomas, and C. Barbieri, 2007: Modeling of the light scattering properties of cometary dust using fractal aggregates, *Astron. Astrophys.*, **461**, 351–364.
3. Borghese, F., P. Denti, G. Toscano, and O.I. Sindoni, 1979: Electromagnetic scattering by a cluster of spheres, *Appl. Opt.*, **18**, 116–120.
4. Borghese, F., P. Denti, R. Saija, G. Toscano, and O.I. Sindoni, 1984a: Multiple electromagnetic scattering from a cluster of spheres. I. Theory, *Aerosol Sci. Technol.*, **4**, 227–235.
5. Borghese, F., P. Denti, R. Saija, G. Toscano, and O.I. Sindoni, 1984b: Use of group theory for the description of electromagnetic scattering from molecular systems, *J. Opt. Soc. Am. A*, **1**, 183–191.
6. Borghese, F., P. Denti, and R. Saija, 1992: Optical properties of spheres containing a spherical eccentric inclusion, *J. Opt. Soc. Am. A*, **9**, 1327–1335.
7. Borghese, F., P. Denti, and R. Saija, 1994: Optical properties of spheres containing several spherical inclusions, *Appl. Opt.*, **33**, 484–493.
8. Bohren, C.F., and D.R. Huffman, 1983: *Absorption and Scattering of Light by Small Particles*, New York: Wiley.
9. Bruning, J.H., and Y.T. Lo, 1969: Multiple scattering by spheres, Antenna Laboratory Rep. 69-5 (Antenna Laboratory, Department of Electrical Engineering, Engineering Experiment Station, University of Illinois, Urbana, Illinois).
10. Bruning, J.H., and Y.T. Lo, 1971a: Multiple scattering of EM waves by spheres, Part I – Multiple expansion and ray-optical solutions, *IEEE Trans. Antennas Propag.*, **AP-19**, 378–390.
11. Bruning, J.H., and Y.T. Lo, 1971b: Multiple scattering of EM waves by spheres, Part II – Numerical and experimental results, *IEEE Trans. Antennas Propag.*, **AP-19**, 391–400.
12. Chylek, P., and V. Srivastava, 1983: Dielectric constant of a composite inhomogeneous medium. *Phys. Rev. B.*, **27**, 5098–5106.
13. Chýlek, P., G. Videen, D.J.W. Geldart, J.S. Dobbie, and H.C.W. Tso, 2000: Effective approximations for heterogeneous particles, in *Light Scattering by Nonspherical Particles: Theory, Measurements, and Geophysical Applications*, M.I. Mishchenko, J.W. Hovenier, J.D. Travis (eds.), San Diego: Academic, 273–308.
14. Collinge, M.J., and B.T. Draine, 2004: Discrete dipole approximation using the surface-corrected lattice dispersion relation, *J. Opt. Soc. Am., A*, **21**, 2023–2028.
15. Dorschner, J., B. Begemann, T. Henning, C. Jäger, and H. Mutschke, 1995: Steps toward interstellar silicate mineralogy. II Study of Mg glasses of variable composition. *Astron. Astrophys.*, **300**, 503–520.
16. Draine, B.T., 1988: The discrete-dipole approximation and its application to interstellar graphite grains. *Astrophys. J.*, **333**, 848–872.

17. Draine, B.T., and P.J. Flatau, 1994: Discrete dipole approximation for scattering calculations, *J. Opt. Soc. Am.*, **A11**, 1491–1499.
18. Draine, B.T., 2000: The discrete-dipole approximation for light scattering by irregular targets, in *Light Scattering by Nonspherical Particles: Theory, Measurements, and Geophysical Applications*, M.I. Mishchenko, J.W. Hovenier, and J.D. Travis (eds.), San Diego: Academic, 131–145.
19. Draine, B.T., and P.J. Flatau, 2004: User Guide to the Discrete Dipole Approximation Code DDSCAT 6.1, http://arxiv.org/abs/astro-ph/0409262.
20. Fuller, K.A., G.W. Kattawar, R.T. Wang, 1986: Electromagnetic scattering from two dielectric spheres: further comparisons between theory and experiment, *Appl. Opt.*, **25**, 2521–2529.
21. Fuller, F.A., and G.W. Kattawar, 1988a: Consummate solution to the problem of classical electromagnetic scattering by ensembles of spheres. I. Linear chains, *Opt. Lett.*, **13**, 90–92.
22. Fuller, K.A., and G.W. Kattawar, 1988b: Consummate solution to the problem of classical electromagnetic scattering by ensembles of spheres. II. Clusters of arbitrarily configurations, *Opt. Lett.*, **13**, 1063–1065.
23. Fuller, K.A., 1994: Scattering and absorption cross sections of compounded spheres. I. Theory for external aggregation, *J. Opt. Soc. Am. A*, **11**, 3251–3260.
24. Fuller, K.A., 1995: Scattering and absorption cross sections of compounded spheres. III. Spheres containing arbitrarily located spherical inhomogeneities, *Appl. Opt.*, **12**, 893–904.
25. Gérardy, J.M., and M. Ausloos, 1980: Absorption spectrum of clusters of spheres from the general solution of Maxwell's equations. The long-wavelength limit, *Phys. Rev. B*, **22**, 4950–4959.
26. Gérardy, J.M., and M. Ausloos, 1982: Absorption spectrum of clusters of spheres from the general solution of Maxwell's equations. II. Optical properties of aggregated metal spheres, *Phys. Rev. B*, **25**, 4204–4229.
27. Gustafson, B.Å.S., 1996: Microwave analog to light scattering measurements: a modern implementation of a proven method to achieve precise control, *J. Quant. Spectrosc. Rad. Transfer*, **55**, 663–672.
28. Gustafson, B.Å.S., 1999: Microwave analog to light scattering measurements, in *Light Scattering by Nonspherical Particles: Theory, Measurements, and Geophysical Applications*, M. I. Mishchenko, J. W. Hovenier, and L. D. Travis (eds.), Academic Press, 367–390.
29. Gutkowicz-Krusin, D., and B.T. Draine, 2004: Propagation of Electromagnetic Waves on a Rectangular Lattice of Polarizable Points, http://arxiv.org/abs/astro-ph/0403082
30. Hanner, M.S., and J.P. Bradley, 2004: Composition and mineralogy of cometary dust, in: Festou, M., Keller, U., and Weaver, H. (eds.), *Comets II*, University of Arizona Press, Tucson, AZ, 555–564.
31. Hapke, B., 1993: *Theory of Reflectance and Emittance Spectroscopy*, New York: Cambridge University Press.
32. Jones, R.V., L. Spitzer Jr, 1967: Magnetic alignment of interstellar grains, *Astrophys. J.*, **147**, 943–964.
33. Jones, A.R., 1979: Electromagnetic wave scattering by assemblies of particles in the Rayleigh approximation. *Proc. R. Soc. Lond. A.*, **366**, 111–127.
34. Kimura, H., 2001: Light-scattering properties of fractal aggregates: numerical calculations by a superposition technique and the discrete-dipole approximation, *J. Quant. Spectrosc. Rad. Transfer*, **70**, 581–594.

35. Köhler, M., H. Kimura, and I. Mann, 2006: Applicability of the discrete-dipole approximation to light-scattering simulations of large cosmic dust aggregates, *Astron. Astrophys.*, **448**, 395–399.
36. Kozasa, T., J. Blum, and T. Mukai, 1992: Optical properties of dust aggregates, I. Wavelength dependence. *Astron. Astrophys.*, **263**, 423–432.
37. Kozasa, T., J. Blum, H. Okamoto, and T. Mukai, 1993: Optical properties of dust aggregates. II. Angular dependence of scattered light, *Astron. Astrophys.*, **276**, 278–288.
38. Laor, A., and B.T. Draine, 1993: Spectroscopic constraints on the properties of dust in active galactic nuclei. *Astrophys. J.*, **402**, 441–68.
39. Lasue, J., and A.C. Levasseur-Regourd, 2006: Porous irregular aggregates of submicron sized grains to reproduce cometary dust light scattering observations, *J. Quant. Spectrosc. Rad. Transfer*, **100**, 220–236.
40. Liang, C., and Y.T. Lo, 1967: Scattering by two spheres, *Radio Sci.*, **2**, 1481–1495.
41. Mackowski, D.W., 1991: Analysis of radiative scattering for multiple sphere configurations, *Proc. R. Soc. London Ser. A*, **433**, 599–614.
42. Mackowski, D.W., 1994: Calculation of total cross sections of multiple-sphere clusters. *J. Opt. Soc. Am. A*, **11**, 2851–2861.
43. Mackowski, D.W., and M.I. Mishchenko, 1996: Calculation of the T matrix and the scattering matrix for ensembles of spheres, *J. Opt. Soc. Am. A*, **13**, 2266–2278.
44. Matsumoto, M., and T. Nishimura, 1998: Mersenne Twister: A 623-dimensionally equidistributed uniform pseudorandom number generator, *ACM Trans. on Modeling and Computer Simulation*, **8**, 3–30.
45. McQueen, J., 1967: Some methods for classification and analysis of multivariate observations. In: *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 281–297.
46. Mishchenko, M.I., 1991: Light scattering by randomly oriented axially symmetric particles, *J. Opt. Soc. Amer.*, **8**, 871–882.
47. Mishchenko, M.I., W.J. Wiscombe, J.W. Hovenier, and J.D. Travis, 2000: Overview of scattering by nonspherical particles. In *Light Scattering by Nonspherical Particles: Theory, Measurements, and Geophysical Applications*, M.I. Mishchenko, J.W. Hovenier, and J.D. Travis (eds.), San Diego: Academic, 29–60.
48. Mishchenko, M.I., L.D. Travis, and A.A. Lacis, 2002: *Scattering, Absorption, and Emission of Light by Small Particles*, Cambridge: Cambridge University Press.
49. Mishchenko, M.I., 2006: Scale invariance rule in electromagnetic scattering, *J. Quant. Spectrosc. Rad. Transfer*, **101**, 411–415.
50. Mishchenko, M.I., L.D. Travis, and A.A. Lacis, 2006: *Multiple Scattering of Light by Particles: Radiative Transfer and Coherent Backscattering*, Cambridge: Cambridge University Press.
51. Mishchenko, M.I., G. Videen, V.A. Babenko, N.G. Khlebtsov, and T. Wriedt, 2007: Comprehensive T-matrix reference database: A 2004–06 update, *J. Quant. Spectrosc. Rad. Transfer*, **106**, 304–324.
52. Mishchenko, M.I., G. Videen, N.G. Khlebtsov, T. Wriedt, and N.T. Zakharova, 2008: Comprehensive T-matrix reference database: A 2006–07 update, *J. Quant. Spectrosc. Rad. Transfer*, **109**, 1447–1460.
53. Muinonen, K. 1990: *Light scattering by inhomogeneous media: Backward enhancement and reversal of linear polarization*, Ph.D. thesis, University of Helsinki, Helsinki.
54. Mukai, T., 1989: Cometary dust and interplanetary particles. In: Bonetti A, Greenberg JM, and Aiello SA (eds.), *Evolution of Interstellar Dust and Related Topics*. Amsterdam: Elsevier, p. 397.

55. Mukai, T., H. Ishimoto, T. Kozasa, J. Blum, and J.M. Greenberg, 1992: Radiation pressure forces of fluffy porous grains. *Astron. Astrophys.*, **262**, 315–320.
56. Mukai, T., J. Blum, A.M. Nakamura, R.E. Johnson, and O. Havnes, 2001: Physical processes on interplanetary dust, in *Interplanetary Dust*, E. Grün, B. Å. S. Gustafson, S. Dermott, and H. Fechtig (eds.), New York: Springer, 445–508.
57. Niederreiter, H., 1992: *Random Number Generation and Quasi-Monte Carlo Methods.* Philadelphia: SIAM.
58. Okada, Y., T. Mukai, I. Mann, H. Nomura, T. Takeuchi, I. Sano, and S. Mukai, 2007: Grouping and adding method for calculating light scattering by large fluffy aggregates, *J. Quant. Spectrosc. Rad. Transfer*, **108**, 65–80.
59. Okada, Y., 2008: Efficient numerical orientation averaging of light scattering properties with a quasi-Monte-Carlo method, *J. Quant. Spectrosc. Rad. Transfer*, **109**, 1719–1742.
60. Okada, Y., I. Mann, T. Mukai, and M. Köhler, 2008: Extended calculation of polarization and intensity of fractal aggregates based on rigorous method for light scattering simulations with numerical orientation averaging, *J. Quant. Spectrosc. Rad. Transfer*, **109**, 2613–2627.
61. Okada, Y., A.A. Kokhanovsky, 2009: Light scattering and absorption by densely packed groups of spherical particles, *J. Quant. Spectrosc. Rad. Transfer*, **110**, doi:10.1016/j.jqsrt.2008.12.007.
62. Petrov, D.V., E.S. Zubko, and Y.G. Shkuratov, 2004: Modeling polarization properties of structure analogs of cometary dust particles, *Proceedings of the 67th Annual Meeting of the Meteoritical Society, August 2–6, Rio de Janeiro, Brazil*, abstract no. 5141.
63. Petrova, E.V., K. Jockers, and N.N. Kiselev, 2000: Light scattering by aggregates with sizes comparable to the wavelength: An application to cometary dust, *Icarus*, **148**, 526–536.
64. Petrova, E.V., V.P. Tishkovets, and K. Jockers, 2004: Polarization of light scattered by solar system bodies and the aggregate model of dust particles, *Solar System Res.*, **38**, 309–324.
65. Petrova, E.V., V.P. Tishkovets, and K. Jockers, 2007: Modeling of opposition effects with ensembles of clusters: Interplay of various scattering mechanisms, *Icarus*, **188**, 233–245.
66. Preibisch, T., V. Ossenkopf, H.W. Yorke, and T. Henning, 1993: The influence of ice-coated grains on protostellar spectra, *Astron. Astrophys.*, **279**, 577–88.
67. Seeliger, H. 1887: Zur Theorie Baleuchtung der Grossen Planeten Inbesondere des Saturn. *Abhandl. Bayer. Akad. Wiss. Math. Naturw. Kl. II*, **16**, 405–516.
68. Sun, W., and Q. Fu, 2000: Finite-difference time-domain solution of light scattering by dielectric particles with large complex refractive indices, *Appl. Opt.*, **39** 5569–5578.
69. Taflove, A., and S.C. Hagness, 2005: *Computational Electrodynamics: The Finite-Difference Time-Domain Method* (Artech House Antennas and Propagation Library). Massachusetts: Artech House.
70. Tamanai, A., H. Mutschke, J. Blum, and R. Neuhäuser, 2006: Experimental infrared spectroscopic measurement of light extinction for agglomerate dust grains, *J. Quant. Spectrosc. Rad. Transfer*, **100**, 373–381.
71. Tishkovets, V.P., E.V. Petrova, and K. Jockers, 2004: Optical properties of aggregate particles comparable in size to the wavelength, *J. Quant. Spectrosc. Rad. Transfer*, **86**, 241–255.
72. Volten, H., O. Muñoz, J.W. Hovenier, F.J.M. Rietmeijer, J.A. Nuth, L.B.F.M. Waters, and W.J. van der Zande, 2007: Experimental light scattering by fluffy aggregates of

magnesiosilica, ferrosilica, and alumina cosmic dust analogs, *Astron. Astrophys.*, **470**, 377–386.
73. Voshchinnikov, N.V., V.B. Il'in, and T. Henning, 2005: Modeling the optical properties of composite and porous interstellar grains, *Astron. Astrophys.*, **429**, 371–381.
74. Wang, Y.M., and W.C. Chew, 1993: A recursive T-matrix approach for the solution of electromagnetic scattering by many spheres, *IEEE Trans. Antennas Propag.*, **41**, 1633–1639.
75. Waterman, P.C., 1965: Matrix formulation of electromagnetic scattering, *Prof. IEEE*, **53**, 805–812.
76. Waterman, P.C., 1971: Symmetry, unitarity, and geometry in electromagnetic scattering. *Phys. Rev.*, **D 3**, 825–839.
77. Wolff, M.J., G.C. Clayton, P.G. Martin, and R.E. Schulte-Ladbeck, 1994: Modeling composite and fluffy grains: The effects of porosity. *Astrophys. J.*, **423**, 412–425.
78. Wolff, M.J., G.C. Clayton, and S.J. Gibson, 1998: Modeling composite and fluffy grains. II. Porosity and phase functions. *Astrophys. J.*, **503**, 815–830.
79. Xu, Y.L., 1995: Electromagnetic scattering by an aggregate of spheres. *Appl. Opt.*, **34**, 4573–4588.
80. Xu, Y.L., 1997: Electromagnetic scattering by an aggregate of spheres: far field. *Appl. Opt.*, **36**, 9496–9508.
81. Xu, Y.L., and N.G. Khlebtsov, 2003: Orientation-averaged radiative properties of an arbitrary configuration of scatterers, *J. Quant. Spectrosc. Rad. Transfer*, **79-80**, 1121–1137.
82. Xu, Y.L., 2004: Radiative-scattering signatures of an ensemble of nonspherical particles, *J. Quant. Spectrosc. Rad. Transfer*, **89**, 385–419.
83. Yang, P., and K.N. Liou, 2000: Finite difference time domain method for light scattering by nonspherical and inhomogeneous particles, in *Light Scattering by Nonspherical Particles: Theory, Measurements, and Geophysical Applications*, M.I. Mishchenko, J.W. Hovenier, and J.D. Travis (eds.), San Diego: Academic, 173–221.
84. Yurkin, M.A., V.P. Maltsev, and A.G. Hoekstra, 2007: The discrete dipole approximation for simulation of light scattering by particles much larger than the wavelength, *J. Quant. Spectrosc. Rad. Transfer*, **106**, 546–557.
85. Zubko, E., Y. Shkuratov, H. Matthew, J. Eversole, and G. Videen, 2004: Backscatter of agglomerate particles, *J. Quant. Spectrosc. Rad. Transfer*, **88**, 163–171.
86. Mischenko, M.I., G. Videen, V.A. Babenko, N.G. Khlebtsov, and T. Wriedt, 2004: T-matrix theory of electromagnetic scattering by particles and its applications: a comprehensive reference database, *J. Quant. Spectrosc. Rad. Transfer*, **88**, 357–406.
87. ftp://ftp.eng.auburn.edu/pub/dmckwski/scatcodes
88. http://yasokada.googlepages.com/gam
89. http://yasokada.googlepages.com/qmc
90. http://www.iup.uni-bremen.de/˜alexk, then choose the link 'T-matrix' at the right side menu

2 Application of scattering theories to the characterization of precipitation processes

Sandra Jacquier and Frédéric Gruy

2.1 Introduction

Solid–liquid suspensions are frequently used in industrial processes. These suspensions usually contain aggregates made up of solid primary particles. Many characterization tools of these suspensions are based on light scattering (Mie theory). However, Mie theory (1908) is not always applicable to practical problems since the scatterer must be a homogeneous sphere. The ordinary particle sizers that use this theory do not make it possible to measure non-spherical particle geometrical characteristics. Extensions of the Mie theory for arbitrary shaped particles or particle aggregates are available nowadays (the T-matrix method, the Generalized Multiparticle Mie (GMM)-solution, etc.). But the computing times of the optical properties via these exact theories do not allow for a real-time analysis. This chapter is therefore dedicated to the search for approximate methods for the estimation of aggregate optical properties, particularly their scattering cross-section.

This chapter is split into five main sections. Section 2.2 concerns the aggregation process and, more generally, the precipitation process, to provide a better understanding of the framework of this study. Precipitation is the formation of a solid in a solution during a chemical reaction. The morphology of particles currently observed during precipitation or particle synthesis will be described. The relationship between optics and particle technology will be recalled. Section 2.3 outlines briefly the different approximate methods used for the case of spherical and non-spherical particles. Section 2.4 presents selected approximations for light scattering cross-sections in the case of aggregates. The quality of each approximation will be discussed by comparison with the exact theory. Practical cases will be presented. Section 2.5 is an extension of the previous section to the light pressure cross-section. Section 2.6 is an attempt to relate the scattering properties of aggregates to their geometrical characteristics.

In the next section we describe the context of the need for approximate methods for scattering cross-section of aggregates.

2.2 Aggregate formation

Materials are often made from inorganic particles. These are formed by reactions in the gas phase or, more commonly, in liquid phases. The main process is called precipitation.

2.2.1 Precipitation and particle synthesis

The classical situation is the following: a solution consisting of a solute A and a solvent is mixed with another one consisting of a solute B and the same solvent. The two solutes react to form a solid product denoted AB.

$$\mathrm{A} + \mathrm{B} \leftrightarrow \mathrm{AB}\,.$$

A and B are often ions.

If equilibrium between the solid phase and the solutes is reached, thermodynamics tells us that this equilibrium (saturated solution) is characterized by the solubility of the solid phase. However, when the concentrations of A and B are high enough to produce solid particles, the initial solution is supersaturated. So, the ratio of the actual concentration and the equilibrium concentration (solubility), called supersaturation, is the key parameter of the dynamics of precipitation. The higher the supersaturation, the higher the precipitation rate. Particle formation can be distinguished into three stages: nucleation, growth and agglomeration.

Nucleation is the birth of particles that are large enough to grow (nuclei). Particles that are too small are unstable and dissolve in the solution. The nucleation rate, J_{N}, is defined as the number of created nuclei per unit volume and per unit time. The formed nuclei are characterized by a so-called critical size. Typical critical size values are within the range 1–50 nm. The nucleation rate is a nonlinear increasing function of supersaturation. The nucleation stage in a precipitation process is often difficult to observe. If light is used as a probe to study precipitation, nuclei can be treated as Rayleigh scatterers. Considering the nuclei concentration, interaction between light and suspension is usually below the detector threshold. Other methods (for instance, small-angle X-ray scattering) are preferred, but are more difficult to apply. Details about the nuclei are not known. Hence, nuclei leading to inorganic particles are often considered as amorphous and spherical.

The following stage is the growth of the nuclei. They may additionally convert to crystals. Crystals present facets, the occurrence of which can be explained from crystallography and thermodynamics. However, depending on the precipitation reaction, only amorphous particles (i.e., hydroxylated compounds) can be found. Usually, the growth rate is defined as the derivative of crystal characteristic length (e.g., the diameter) with time. The growth rate increases with the supersaturation. Most often the relation between growth rate and supersaturation is linear.

For different reasons (one is the decrease of supersaturation during precipitation, due to the mass transfer from solution to crystals), an upper size limit for the growing crystals is observed. At the end of the second stage, crystal or particle size ranges between 0.1 and 10 μm. During this period, particles become large enough to scatter light effectively. Thus, light scattering methods are suitable for studying the growth of crystals or primary particles in suspension.

Before or after reaching the end of growth, crystals or amorphous particles can form clusters or agglomerates. Agglomeration requires the collision of particles and their subsequent adhesion due to attractive forces (for instance, van der Waals forces). Consolidation between primary particles or crystals can take place by crystalline growth from the contact point. In quiescent liquids, the particle collisions are due to their Brownian motion. For flowing suspensions, collisions are due to Brownian motion for small particles or crystals (smaller than about $0.2\,\mu$m) and to local shear for larger particles.

When the agglomerate increases in size, it becomes fragile. Then, break-up takes place and a limit size is reached (with values in the range 5–$100\,\mu$m). The stress acting on the agglomerate results from the shear, but also the transition between viscous and inertial turbulent regimes. This often leads to agglomerates containing only a few primary particles (less than one hundred primary particles).

Summarizing: inorganic solutes can lead to solid particles in a batch precipitation vessel, provided that the solution is initially supersaturated. Nucleation, i.e. the birth of nuclei with a critical size, takes place, while supersaturation decreases as a consequence. Finally, supersaturation becomes too small, to produce new nuclei. Thus, there is a mass transfer from solution to the nuclei surface leading to the growth of nuclei. In this way, crystals or amorphous particles are formed. As supersaturation tends to one, growth stops. Depending on the surface charge of particles, their agglomeration may occur throughout the growth period.

2.2.2 Particle shapes during precipitation

Images from electron microscopy often show the complexity of particle structure. Particles formed by growth from solution can be crystals with well-defined crystalline faces, but may be agglomerates of smaller (nanometric) particles. In this case, agglomeration can be due to Brownian motion and the subsequent collision of particles, but also to the contact of specific crystalline faces belonging to two particles. The first situation leads to random agglomerate with spherical symmetry. The second phenomenon, also called orientated agglomeration, leads to regularly shaped particles (i.e., cylinders as disks stack). It must be underlined that the mechanism of orientated agglomeration is still being studied. Whatever the structure of particle, crystals, random or ordered agglomerate of nanoparticles, their geometric characteristics are easily determined by image analysis of electron micrographs.

However, these particles are rarely separate. They form agglomerates after collision due to Brownian motion and/or shear flow. The structure of agglomerates is disordered and is often considered as fractal-like. However, the reader must keep in mind that these agglomerates consist of only a few particles.

Fig. 2.1 presents agglomerates of $SrMoO_4$, strontium molybdate, crystals (Cameirao et al., 2008). They are obtained by precipitation:

$$SrCl_2 + Na_2MoO_4 \rightarrow SrMoO_4 + 2NaCl\,.$$

Bipyramidal crystals in the size range 3–$10\,\mu$m are formed. Agglomerate size is in the range 20–$80\,\mu$m.

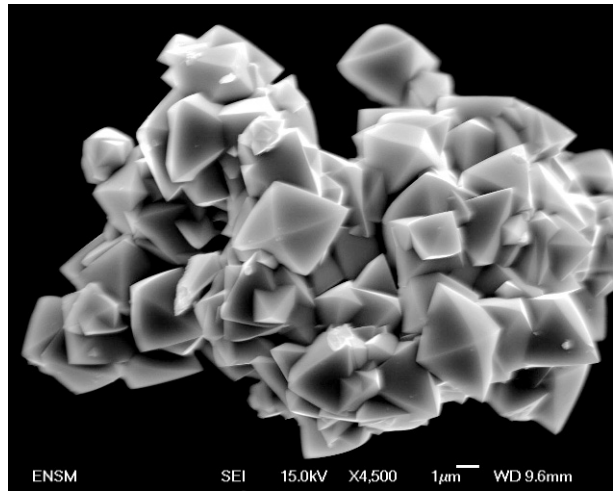


Fig. 2.1. Agglomerate of $SrMoO_4$ crystals.

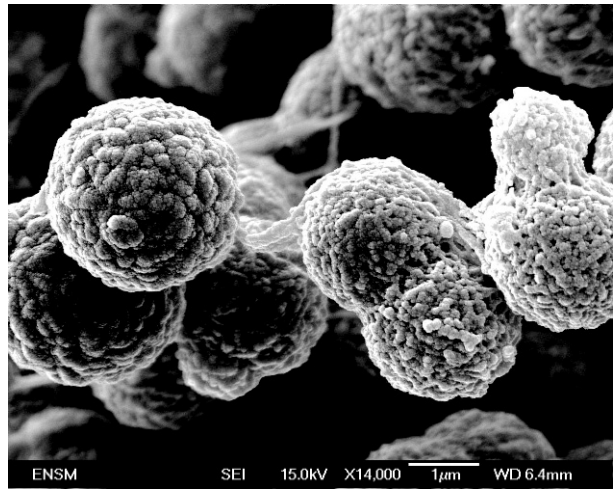


Fig. 2.2. Multi-scale ZnS agglomerate.

Fig. 2.2 presents agglomerates of zinc sulphide particles (Mekki-Berrada et al., 2005). They are obtained by homogeneous precipitation:

$$ZnSO_4 + CH_3CSNH_2 + 2H_2O \xrightarrow[80°C]{H^+} ZnS + CH_3CO_2H + NH_4^+HSO_4^- \,.$$

ZnS particles are sphere-like with a mean size equal to $3\,\mu$m. They consist of nanoparticles, 30 nm sized (one may observe an intermediate structure in the range $100 \ldots 300$ nm). Micro-particles seem relatively dense. However, porosity and inner structure depend on the acidity and temperature. Micro-particles collide to form agglomerates in the range 20–60 μm.

Fig. 2.3 shows ordered agglomerates resulting from the stacking of $Ni(OH)_2$ nanosized disks (Coudun et al., 2007). They are obtained by precipitation from nickel di-dodecylsulfate and ammonia:

$$Ni(DS)_2 + 2OH^- \xrightarrow[60°C]{NH_3} Ni(OH)_2 + 2DS^- .$$

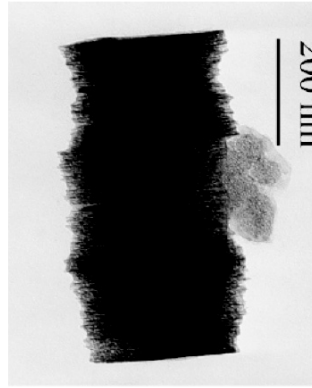


Fig. 2.3. $Ni(OH)_2$ nanosized agglomerates.

2.2.3 Dynamics of precipitation: modelling

In order to manage the complexity of precipitation dynamics, each particle in the precipitation reactor is characterized by space coordinates (x, y, z) and internal coordinates p_i (i.e. radius, characteristic lengths of crystal, volume, porosity, gyration radius, fractal dimension, ...) with $1 \leq i \leq P$.

The larger the number P, the more comprehensive the description of the particles. So, the population density function $n(x, y, z, p_1 \dots p_P)$ is such that $dN = n\, dx\, dy\, dz\, dp_1 \dots dp_P$ represents the number of particles with x in the range $[x; x + dx]$, p_i in the range $[p_i p_i + dp_i]$.

This population density obeys the population balance equation (PBE), that can be formally written as:

$$\frac{\partial n}{\partial t} = -\vec{\nabla} \cdot (\vec{v} n) - \sum_{i=1}^{P} \frac{\partial}{\partial p_i}(G_i(S)n) + J_N(S)\delta(p_1 - p_{1c}) \dots \delta(p_P - p_{Pc}) + V_{ag} \,, \quad (2.1)$$

where $\vec{v}$ is the particle velocity, t is the time, G_i is the growth rate for the p_i parameter $[G_i = (dp_i/dt)]$, δ is the Dirac function, p_{ic} is the internal parameter corresponding to the critical nucleus, $J_N(S)$ is the nucleation rate as a function of the supersaturation S, and V_{ag} is the agglomeration rate.

More often, the following assumptions are used:

- homogeneous suspension
- only one internal parameter (particle radius)

– agglomeration taken into account only at the end of nucleation and growth
– fractal-like agglomerates with constant fractal dimension

For fractal-like agglomerates, the relation between the number j of primary particles inside the agglomerate and its outer radius R_j is:

$$R_j = a \left(\frac{j}{S_r} \right)^{1/D_F} , \tag{2.2}$$

where a is the radius of the primary particle. D_F and S_r are respectively the fractal dimension and the structure factor, which is a function of D_F.

Attempts to consider two internal parameters (radius and length for particles such as a needle, solid volume and surface area for porous particles) have also been made (Tandon and Rosner, 1999; Kostoglou et al., 2006).

Summing up, precipitated particles have a multi-scale structure. One commonly observes three characteristic length values: 20 nm, 2 μm, 20 μm. Sometimes, only two (2 μm, 20 μm) are observed. The smallest particles are most often dense and spherical. The intermediate particles are relatively dense and have a well-defined shape (i.e. sphere, cylinder, ellipsoid, etc.). The largest scale corresponds to disordered or random agglomerates. The reader interested in details of precipitation and population balance may refer to Sugimoto (2000) and Randolph and Larson (1988).

2.2.4 Particle sizing during precipitation

Considerable efforts are made to understand the precipitation mechanisms and to predict the change of the particle morphology with time. On the other hand, industrialists need to monitor and control the precipitation process. Off-line size measurements (i.e. using filtration, powder drying and scanning electron microscope (SEM) observations or suspension sampling and sizing with granulometers) are now avoided because these operations modify the particle morphology. On-line measurements (i.e., using a recirculation loop with a measurement cell in a granulometer) are possible, but representative sampling is difficult to carry out. In-line measurements are preferred. They are often based on light extinction and are obtained from turbidimetric sensors. So, turbidimetry will be at the centre of this chapter.

Typically, the optical sensor for particle sizing is not the only one in the precipitating suspension. The temperature probe and concentration sensors are always introduced into it. Thus, supersaturation and solid fraction (via mass balancing from solute concentration) are deduced and, as a consequence, are known.

A turbidimetric sensor records the transmitted light intensity I_t. One defines the extinction coefficient or turbidity[1] as (see, for instance, Elimelech et al., 1995):

$$\tau = -\frac{1}{L} \ln \frac{I_t}{I_0} . \tag{2.3}$$

[1] Definitions and notations in Eqs. (2.3)–(2.5) are used by researchers in the field of particle sizing concerning suspensions. Other researchers prefer these ones: $\tau^* = -\ln(I_t/I_0)$ and $\sigma_{ext}(\lambda) = \tau^*(\lambda)/L$ where τ^* and σ_{ext} are the optical thickness and the extinction coefficient, respectively.

I_0 and L are the incident light intensity and the geometrical thickness of the medium, respectively. For a monodisperse diluted suspension, the turbidity obeys the equation:

$$\tau(\lambda) = N(x_s, y_s, z_s, p_1 \dots p_P)\, C_{ext}(\lambda, p_1 \dots p_P) \tag{2.4}$$

The functions N and C_{ext} are the particle number concentration and the extinction cross-section of the particles. (x_s, y_s, z_s) are the coordinates of the sensor in the reactor.

For a complex diluted suspension, turbidity contains the contribution of each kind of particles:

$$\tau(\lambda) = \int_{[p_1 \dots p_P]} n(x_s, y_s, z_s, p_1 \dots p_P)\, C_{ext}(\lambda_i p_1 \dots p_P)\, dp_1 \dots dp_P\,. \tag{2.5}$$

Thus, the turbidity monitoring gives information on the population density change with time during the precipitation process.

Two strategies can be considered for the analysis of turbidity signals:

- inversion of the integral equation (Eq. (2.5)) in order to get the population density. Then, comparison to PBE (Eq. (2.1)) solution and identification of unknown physicochemical and morphological parameters (i.e., fractal dimension ...)
- PBE (Eq. (2.1)) solving; calculation of $\tau(\lambda)$ (Eq. (2.5)); comparison with experimental turbidity for deducing the unknown physicochemical and morphological parameters

For numerical reasons, the last one is preferred. However, whatever the strategy, the knowledge of $C_{ext}(\lambda, p_1 \dots p_P)$ is needed.

The extinction cross-section is dependent on the relative refractive index m, which is the ratio between the refractive indices of the material and the suspending medium. Three typical cases can be envisaged: low optical contrast m–1 value ($0 < m - 1 < 0.1$), moderate contrast m–1 value ($0.1 < m$–$1 < 0.5$) and high optical contrast m–1 value (m–$1 > 0.5$). Corresponding materials (suspended in water) could be, respectively, silica SiO_2 ($m = 1.08$), alumina Al_2O_3 ($m = 1.20$) and titania TiO_2 ($m = 2$).

We will focus our work on non-absorbing (in the visible range) materials that are most commonly found in the precipitation process. Thus, scattering C_{sca} and extinction C_{ext} cross-sections are equal.

It is obvious that PBE solving, cross-section calculation and optimization procedure require great computational efforts and make it difficult to get results in a short time. Therefore, any rapid calculation of the optical cross-sections would be a step forward. The accuracy of approximations has to fit the measurement accuracy. In the case of turbidity, measurements within 3% error can be considered as satisfactory.

The need for approximations is particularly important for agglomerates[2]. Fast calculations have to be based on known approximations coming from light scattering theory. The next section briefly recalls them.

2.3 Approximations for non-spherical particles

The scattering cross-section is a function of the dimensionless particle size parameter x ($= ka$ for a sphere), the particle and the medium optical refractive indices respectively denoted n_p and n_m, the wavelength λ (and the wave number $k = 2\pi/\lambda$) of the incident light in the medium and the orientation of the incident light in the relation of the particle position. The relative refractive index m, which is used in the following equations, is the ratio between the material refractive index n_p and the medium refractive index n_m.

The exact theory was developed for a sphere in 1908 by G. Mie (see van de Hulst, 1981) and for spheroids by several authors (Asano and Yamamoto, 1975; Asano, 1979; Asano and Sato, 1980; Voshchinnikov and Farafonov, 1992; Farafonov et al., 1999).

In this section, three classical approximations are recalled: the Rayleigh approximation, the Rayleigh–Gans–Debye approximation and the Anomalous Diffraction approximation. Principles are presented and an application is given for a sphere. The reader interested in details on scattering theories may refer to Van de Hulst (1981) and Kokhanovsky (2001).

2.3.1 Rayleigh approximation

The Rayleigh approximation that considers the scatterer as an oscillating dipole has a validity range of $x \ll 1$, $|mx| \ll 1$. So, the scattering efficiency factor for a sphere is:

$$Q_{sca} = \frac{8}{3} x^4 \left| \frac{m^2 - 1}{m^2 + 2} \right|^2 \tag{2.6}$$

and the scattering cross-section is $C_{sca} = Q_{sca} G$ (G represents the particle projected area, for a sphere $G = \pi a^2$).

A comparison between this approximation and the Mie exact theory shows that the validity range, in terms of maximum size, varies according to the relative refractive index and the scattering angle (Mishchenko et al., 2002, 2004).

2.3.2 Rayleigh–Gans–Debye approximation

The validity range of the Rayleigh–Gans–Debye approximation (RGD) is: $2x|m - 1| \ll 1$ and $|m - 1| \ll 1$.

[2]The term aggregation corresponds to the formation of a cluster, the primary particles of which only interact by physical forces such as van der Waals forces. On the other hand, agglomeration is aggregation followed by strengthening at the contact point in a supersaturated solution. Aggregate and agglomerate optical properties will be treated in the same way.

Fig. 2.4 represents a particle with an unspecified shape lit by a plane wave being propagated along the axis z'. It is supposed that each volume element is a Rayleigh scatterer and behaves independently of the other particle volume elements. The scattering waves of all these volume elements interfere. The phases of all these scattering waves are ascribed to a common coordinates reference in order to handle their amplitude.

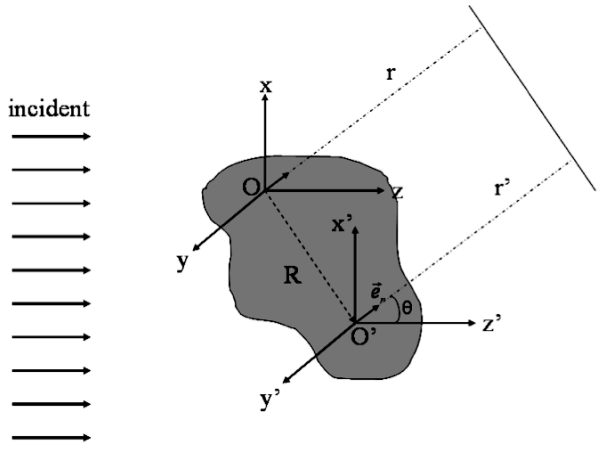


Fig. 2.4. RGD approximation.

The expression of the contribution, of the volume element ΔV located out of O, to the scattering field by the particle is:

$$\begin{pmatrix} \Delta E_{\parallel,sca} \\ \Delta E_{\perp,sca} \end{pmatrix} = \begin{pmatrix} S_2 & 0 \\ 0 & S_1 \end{pmatrix} \frac{e^{i\cdot k\cdot r - i\cdot k\cdot z}}{-i\cdot k\cdot r} \Delta V \begin{pmatrix} E_{\parallel,inc} \\ E_{\perp,inc} \end{pmatrix} .$$

The contribution of a volume element located in O' will be:

$$\begin{pmatrix} \Delta E_{\parallel,sca} \\ \Delta E_{\perp,sca} \end{pmatrix} = \begin{pmatrix} S_2 & 0 \\ 0 & S_1 \end{pmatrix} \frac{e^{i\cdot k(r-z)}}{-i\cdot k\cdot r} \Delta V\, e^{i\delta} \begin{pmatrix} E_{\parallel,inc} \\ E_{\perp,inc} \end{pmatrix} .$$

with $\delta = k\vec{R} \bullet (\vec{e}_z - \vec{e}_r)$ and $\vec{R} = \overrightarrow{OO'}$. S_1 and S_2 are the amplitude functions per volume unity:

$$S_1 \approx -\frac{ik^3}{2\pi}(m-1) \text{ and } S_2 \approx -\frac{ik^3}{2\pi}(m-1)\cos\theta\,.$$

θ is the scattering angle.

Integration is done with respect to particle volume V to obtain the total field in the direction $\vec{e}_r$. So, the amplitude functions for the particle are:

$$S_1 = -\frac{ik^3}{2\pi}(m-1)Vf \text{ and } S_2 = -\frac{ik^3}{2\pi}(m-1)Vf\cos\theta\,.$$

The form factor f obeys:

$$f = \frac{1}{V}\int_V e^{i\delta}\, dV\,, \tag{2.7}$$

that becomes for a sphere:

$$f(u) = \frac{3}{u^3}(\sin u - u\,\cos u),\;\; u = 2x\,\sin\frac{\theta}{2}\,.$$

It follows for a spherical particle (Van de Hulst, 1981) that:

$$Q_{sca} = \int_0^{\pi} F(\theta)\, d/\theta(\pi a^2)\,, \tag{2.8a}$$

where,

$$F(\theta) = \pi a^2 \frac{4}{9}|m-1|^2 x^4 f^2\left(2x\,\sin\frac{\theta}{2}\right)(1+\cos^2\theta)\,\sin\theta\,. \tag{2.8b}$$

2.3.3 Anomalous Diffraction approximation

This approximation, due to Van de Hulst, bears the name of anomalous diffraction (AD) because, for low optical contrast, the light passing through the particle (transmitted without deflection) interferes with that diffracted, then producing a diffraction known as anomalous (Fig. 2.5).

Let us consider particles such as: $x \gg 1$ and $|m_r - 1| \ll 1$ (see the discussion of Videen and Chylek (1998) and Liu (1998)).

The second condition implies that the rays are not deviated when they cross the interface particle-medium and that the reflection is negligible with the same interface. Extinction is therefore due to:

- absorption of the light passing through the particles
- interferences between the light passing through the particle and that passing around it

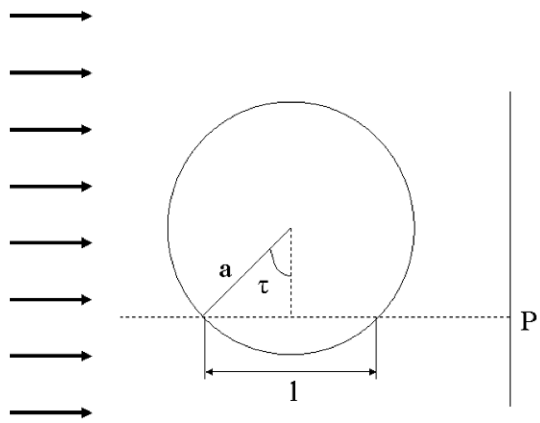


Fig. 2.5. Anomalous Diffraction approximation; *ray passing through a sphere.*

Following Van de Hulst (1981), one derives:

$$C_{ext} = \frac{4\pi}{k^2} \operatorname{Re}\{S(0)\} \text{ with } S(0) = \frac{k^2}{2\pi} \iint_{[S_p]} (1 - e^{-ik(m-1)l})\, dy\, dz\,.$$

Integration is performed over the particle projected area S_p on a plane perpendicular to the propagation direction. l is the computed path of light through the particle, which is a function of the projection coordinates x and y. The integrand represents the subtraction with 'the part of shade' (value 1) of the rays passing through the particle ($e^{-ik(m-1)l}$).

If m is real,

$$C_{sca} = 2 \iint_{[S_p]} \left[1 - \cos(kl(m-1))\right] dS_p\,. \tag{2.9}$$

Therefore, it follows for a sphere

$$Q_{sca} = Q_{ext} = 2 - \frac{4}{\rho} \sin\rho + \frac{4}{\rho^2}(1 - \cos\rho)\,, \tag{2.10}$$

where $\rho = 2x(m-1)$.

The anomalous diffraction was applied to a sphere and an infinitely long circular cylinder (Van de Hulst, 1981), a prism column (Chylek and Klett, 1991), a hexagonal crystal of ice (Sun and Fu, 1999), ellipsoids (Lopatin and Sid'ko, 1988; Streekstra and Hoekstra, 1994), a short cylinder (Liu et al., 1998) and other various forms (Sun and Fu, 2001; Yang et al., 2004).

A comparison between AD and the exact theory (Liu et al., 1998) suggests that AD estimates the extinction more precisely in the case of a random orientation of non-spherical particles than for spheres.

The next section treats approximations for the case of clusters of spheres. The derived approximations are directly related to the previous ones.

2.4 Approximations for aggregate scattering cross-section

This part begins with a short summary of the exact methods. It is followed by a study of the relation between aggregate scattering cross-section and physical characteristics. Finally, four approximations are described and an illustration in the field of precipitation is presented.

2.4.1 Exact theory for non-spherical particles and aggregates

The presented summary (see Table 2.1) of the different exact methods is not exhaustive. But we try to show several methods used to calculate the optical properties of an aggregate. We invite the reader to consult the article of Kahnert (2003) to have a more complete range of these methods and the papers series of Mishchenko et al. (2007, 2008). It is rather difficult to classify them precisely and especially to enumerate all of them. Nevertheless one can classify them in three main categories:

Table 2.1. Methods for treating the light scattering by a non-spherical particle or an aggregate (N is the number of operations in the algorithm (Kahnert, 2003))

Name	Principle	Applications	Strong and weak points
Methods based on partial derivative equations			
SVM Separation of Variable Method	Method applied in the case of the Mie theory; it can be applied when the boundary of the considered particle coincides with the reference frame.	In any reference frame, where the variables separation can be applied. Asano and Yamamoto (1975) used this technique to determine the optical properties of a spheroid.	– The solution obtained is known as exact but calculations are long – the operation for each orientation must be repeated – $N \approx O(x^3)$
FDTD Finite Differences Time Domain	This method consists of discretizing the Maxwell equations, in space and time. Then to solve them it is necessary to start from the initial values (Yang and Liou, 2000).	All particle shapes.	– the operation for each orientation must be repeated – $N \approx O(x^4)$
FEM Finite Element Method	This method consists in discretizing the Helmholtz equation in space and solving numerically using the boundary conditions (Coccioli et al., 1996).	All particle shapes.	– the operation for each orientation must be repeated – the precision depends on the grid which must be selected according to the particle shapes – $N \approx O(x^7)$
PMM Point-Matching Method	In this method, the internal and external fields are expressed as a spherical harmonic vector. The tangential field at the boundary of a particle must be continuous for a fixed number of points belonging to the particle surface.	Normally all particle shapes, but problems are known for the elongated geometries.	– this method is limited to the quasi-spherical particles, it has a dubious convergence, and thus, requires a long CPU time (Wriedt, 1998).

Table 2.1. (Continued)

Name	Principle	Applications	Strong and weak points
Volume or surface integration			
VIEM Volume Integral Equation Method	The field inside and outside the volume is expressed in terms of incident and internal fields for the selected volume. – the internal field is evaluated by considering, for each volume element, as being constant: MOM (Method of Moments) (Harrington, 1968). Alternatively, – each element is regarded as a dipole: DDA (Discrete Dipole Approximation) (Draine and Flatau, 1994)	Inhomogeneous, anisotropic particles.	– MOM and DDA have a long CPU time – the operation for each orientation must be repeated – $N \approx O(x^9)$

1. Methods based on the partial derivative equations which calculate the scattering field by solving the Maxwell equations or the Helmholtz equation. They are subjected to the boundary conditions suitable in the time or the frequency domain.
2. Methods based on integration over volume or surface of equations derived from the Maxwell equations. Thus, the boundary conditions are automatically included in the solution.
3. The other methods are known as hybrids since they derive from the various approaches.

It is important to specify the meaning of the 'T-matrix method' expression which is found in a lot of publications. In the T-matrix method, the incident and scattering fields are expressed in the form of a series of spherical vector wave functions. This approach is named the T-matrix method when the expansion coefficients of the incident wave and the scattering wave are connected by a linear transformation (T is for transition). This matrix T contains all the information on the particle's optical properties for a given wavelength. It is a function of the size parameter, the shape, and the optical refractive index of the considered particles, but it does not depend on the incident field. Thus this matrix is not to be computed at each particle orientation change or change of the incident wave direction.

To classify the publications set relying on this method, a database review was carried out by Mishchenko et al. (2004, 2007, 2008). This method is, in fact, a technique of calculation, which is found associated with various methods (e.g., SVM). Thus, any method making it possible to formulate the problem in the way of a matrix T is called the T-matrix method.

The solving by separation of variables (SVM) for only one sphere can be enlarged to an aggregate of spheres by using the translation theorem for the spherical wave vector functions which expresses them in various bases of coordinates, and, by applying a superposition principle. The total scattering field for an aggregate is then represented by the superposition of the individual scattering fields resulting from each particle in the knowledge that these fields are interdependent. Moreover, one can formulate the problem as a T-matrix. This method is very precise but its computation time depends on the number and the size parameter of primary particles.

We will use in the continuation of this text a method which is in fact a particular case of the T-matrix method (Mishchenko et al., 2004) bearing the name of GMM (Generalized Multi-particle Mie-solution).

We did not find a comparison of the various methods, except for an article of Hovenier et al. (1996) which compared the T-matrix (method by surface integration), SVM and DDA. This article shows that the last is not completely in agreement with the two other methods. As no study was undertaken in this direction, the work presented in this article is achieved with a method which seems, closest to the one used for a simple sphere and validated by experiment (Xu and Gustafson, 2001): GMM. The details of GMM are given by Xu (1995, 1996, 1997a,b, 1998a).

2.4.2 Main features of the scattering properties of aggregates

We present a summary of the main features of the scattering properties of aggregates. Results come from the calculated optical properties of aggregates such as:

- aggregates of spherical primary particles
- number of primary particles in the aggregate $N \in [1, 100]$
- primary particle size parameter (x) in the range: [0.013, 9.25].
- three different materials (SiO_2, Al_2O_3, TiO_2); non-absorbing materials

Optical properties are calculated by using GMM theory.

First of all it is interesting to study the effect of the inter-particle distance on the scattering cross-section.

2.4.2.1 The case of a two-sphere set

The evolution of the scattering cross-section of a two-sphere set according to the type of material, their size parameter and the center to center distance has already been studied by Mishchenko et al. (1995) and Quirantes et al. (2001).

The K_{Xu} parameter for different materials was used. K_{Xu} is defined as the ratio between the scattering cross-section of an aggregate and the sum of scattering cross-sections of N primary particles which form this aggregate (so, the denominator assumes non-interacting and non-interfering spheres).

$$K_{Xu} = \frac{C_{Xu,N}}{NC_{Mie,1}} \,. \tag{2.11}$$

The two spheres, denoted i and j, were gradually separated (center to center) by a factor F_{ij} proportional to their diameter, until they did not interact any more (K_{Xu} is equal to 1). As the separation distance is denoted d_{ij}, the factor obeys the relation:

$$F_{ij} = d_{ij}/(2a)\,,$$

where a is the radius of the primary particle.

According to Fig. 2.6, for size parameter smaller than 0.5, the smaller the primary particle is, the greater the distance factor must be to obviate any interaction. We therefore endorse the conclusion of Kolokolova and Gustafson (2001): a suspension consisting of Rayleigh scatterers as primary particles has to have a very weak volume fraction to avoid multiple scattering, whatever the relative refractive index.

Interaction between particles cannot be ignored even if $F_{ij} > 4$ (for the whole range of the size parameter).

For spheres in contact, K_{Xu} (Fig. 2.6) increases with decreasing size parameter up to a value close to 2. When the two spheres are large enough, the deviation from the non-interacting limit is negligible (e.g., $K_{Xu} < 1.1$ for $x > 5$).

It therefore appears useful to evaluate the critical inter-particle distance for which interaction is negligible. An approximate method for aggregate optics calculation could take it into account.

2.4.2.2 The case of aggregate ($N > 2$)

Auger et al. (2003) studied the relation between the extinction cross-section of an aggregate, its shape (linear or compact configuration) and the number of primary particles (2, 4, 8, 13) in the case of titania TiO_2 (the optical refractive index being equal to 2.8). In this article, the average extinction cross-section (average on the polarization and the incident wave direction) divided by the aggregate volume (made up of monosized spheres) is calculated, according to their particle radius (between 0.04 μm and 0.132 μm). It is found that there exist two size ranges (for $\lambda = 0.546\,\mu$m). For a primary particle radius smaller than 0.08 μm–0.09 μm, an isolated primary particle scatters less than if it was contained in an aggregate. Primary particles belonging to the second range behave in an opposite way. They show that there exist two size ranges concerning the effect of aggregate shape: in the first range, a compact configuration scatters more than linear configuration, and conversely for the second range. Lastly, a comparison with the equivalent sphere approximation shows that the latter is not suitable. Auger et al. (2005) and Jacquier and Gruy (2007a) perform similar studies in the way that they compare the scattering cross-section for various configurations. The Auger et al. (2005) study is based on the distribution of randomly generated aggregates by classical mechanisms of aggregation.

Jacquier (2006) and Jacquier and Gruy (2007a) enlarged the study using different optical refractive indices, a greater range of the primary particle size parameter, and different configurations. They noted the effect of the primary particle number and the aggregate morphology. The results are summarized in the two next paragraphs.

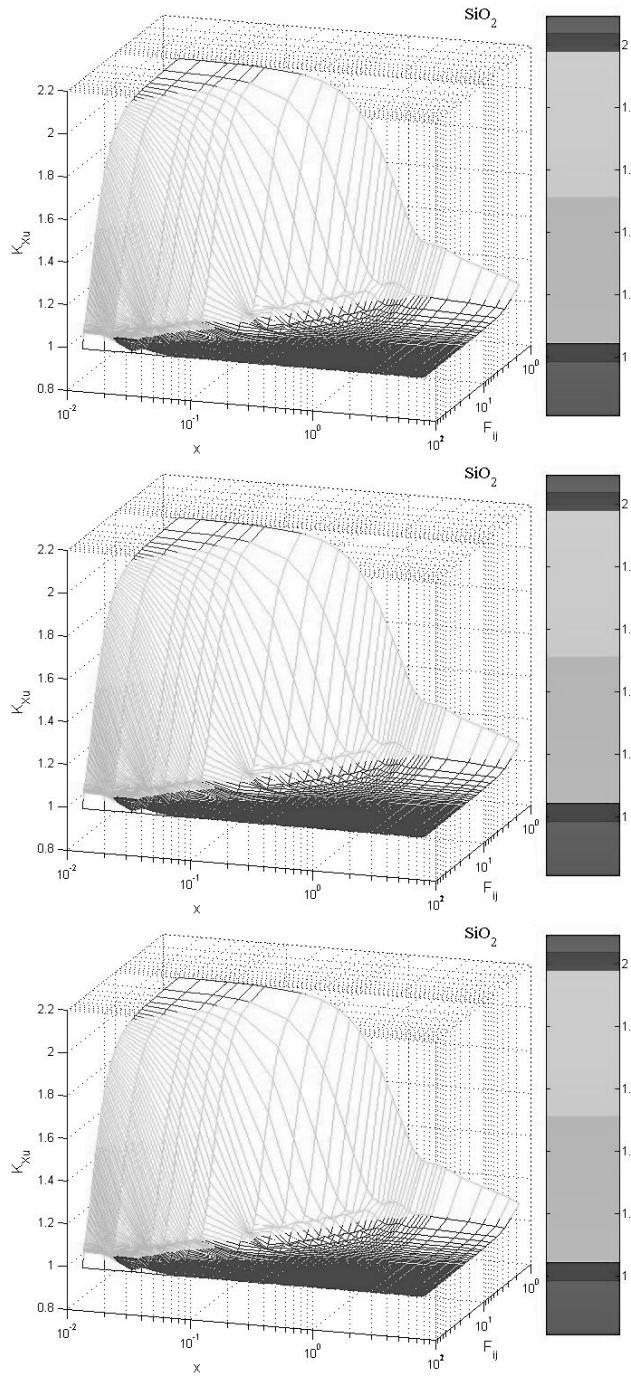


Fig. 2.6. Evolution of K_{Xu} according to the size parameter and the distance factor ranging between 1 and 100 for the three materials (SiO_2, Al_2O_3, TiO_2).

• *Effect of the number of primary particles on the scattering cross-section*
In the paper of Jacquier and Gruy (2007a), two extreme configurations were compared (linear and compact configuration). For each configuration, Jacquier and Gruy noticed that there exist two ranges. The first is for K_{Xu} larger than 2, and the second is for K_{Xu} ranging between 0 and 2. The value of the size parameter of the primary particles corresponding to the range boundary is a weak function of the optical refractive index, the primary particle number and the configuration. However, the authors suggest the first range for $x \in [0, 2]$ (Fig. 2.7(a)) and the second one for $x \in [2, 10]$ (Fig. 2.7(b)) (the limit of their study is for a size parameter smaller than 10). They conclude (as shown in Fig. 2.7):

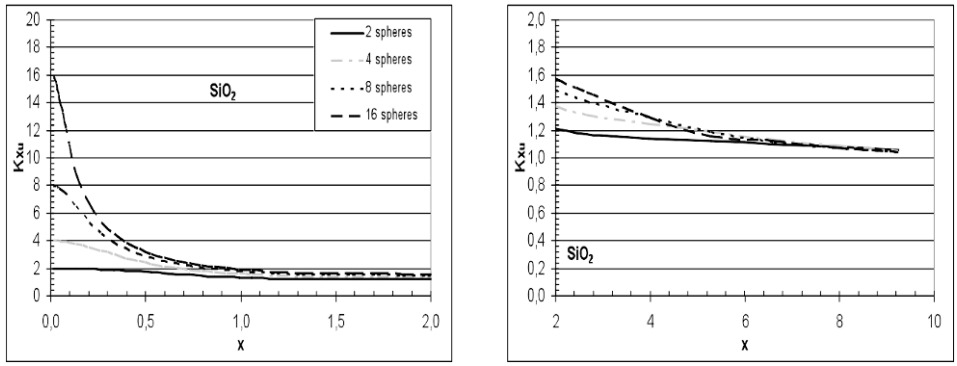
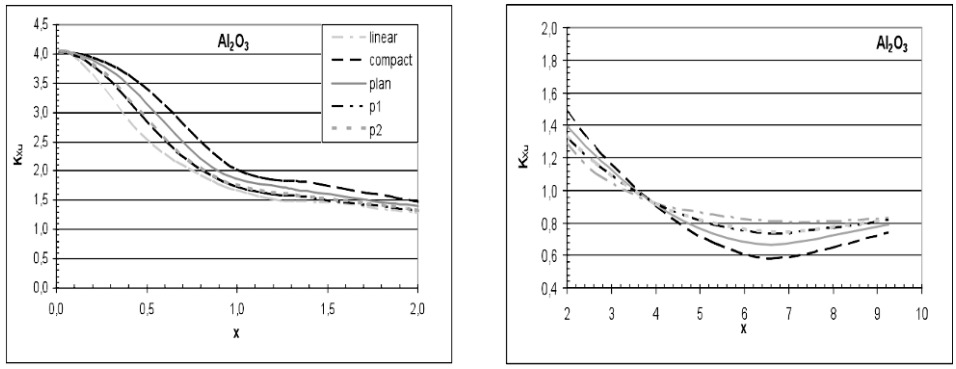


Fig. 2.7. K_{Xu} as a function of the size parameter for the linear configuration with N primary particles (2, 4, 8, 16) and for SiO_2 (Jacquier and Gruy, 2007a): (a) for the range $x \in [0, 2]$ and (b) $x \in [2, 10]$.

– for $x \in [0, 2]$, the larger the number of primary particles, the larger the scattering cross-section whatever the configuration. Indeed, in the case of very small size parameter, the aggregate scattering cross-section is proportional to the particle number square and to the primary particle scattering cross-section ($C_{Xu,N} \propto N^2 C_{Mie,1}$). This relation is checked on a lesser scale by aggregates with high refractive index (e.g., TiO_2). In addition the decrease of $K_{Xu}(x)$ seems to depend on the configuration.
– for $x \in [2, 10]$, $K_{Xu}(x)$ is not yet equal to 1 (Fig. 2.7(b)), i.e. the aggregate scattering cross-section is not the sum of the scattering cross-sections of each primary particle.

• *Effect of the aggregate morphology on scattering cross-section*
As illustrated in Fig. 2.8, it is possible to establish a classification of the configurations according to their scattering cross-section. In the x-domain, where $K_{Xu} > 1$, the scattering cross-section of the compact configuration is higher than that of the plane configuration, itself higher than that of p1 and p2 configurations (which are



a b

Fig. 2.8. K_{Xu} as a function of the size parameter for aggregates with four primary particles (linear, compact, plan, p1, p2) for Al_2O_3.

very close Fig. 2.9). The linear configuration is the weakest. The order is reversed for the other x-domain ($K_{Xu} < 1$). Thus, there are two extreme configurations, linear and compact, between which are other configurations.

compact linear plan p1 p2

Fig. 2.9. Different aggregate configurations in the case of four primary particles.

The primary particles arrangement, i.e. the aggregate configuration, is not without effect on the scattering cross-section, nevertheless, the number of primary particles in it is the prevailing parameter in the range $x \in [0; 2]$. In the second range ($x \in [2; 10]$), the configuration is more important than the number of primary particles.

2.4.2.3 Conclusion on aggregate scattering cross-section

The study of scattering (C_{sca}) cross-sections of aggregates obtained with the exact method revealed that:

- the distance between particles is a relevant parameter for C_{sca},
- different aggregate configurations, following its shape or the number of primary particles which it contains, are perfectly distinguishable,
- the number of primary particles is the relevant parameter in the case of small size parameter x ($x \to 0$, $C_{Xu,N} \propto N^2 C_{Mie,1}$),

– there exists, for an aggregate with a given number of primary particles, two extreme configurations (linear and compact) between which the cross-sections of the others are located.

2.4.3 Approximate methods (CS, BPK, AD, ERI) for aggregates

In this subsection, we describe different approximate methods: the Compact Sphere method (CS), the Berry–Percival–Khlebtsov method (BPK), the Anomalous Diffraction method (AD), and the Effective Refractive Index method (ERI). A first comparison between these four methods was published by Gruy (2001) in connection with aggregation of SiO_2 micro-particles in water.

The study of the parameters influencing the optical properties of aggregate began with Fuller and Kattawar (1988a, b). Rouleau (1996) compared several approximate methods for optical properties based on:

– RGD approximation
– Non-interacting spheres
– Equivalent volume sphere
– Equivalent projected area sphere

This study was carried out only for compact aggregates with 30 primary particles whose size parameter was smaller than 0.6 and the refractive index was kept constant ($m = 1.7 + 0.1i$). He concluded that the abovementioned methods are not efficient except the one using the projected area.

The differences between the methods quoted in the next paragraphs are evaluated in Table 2.2. We chose to differentiate porosity and arrangement. The validity range of all these methods is normally the whole size parameter range except for AD, which, as already mentioned in subsection 2.3.3, is to be used only in the case of large particles.

Table 2.2. Comparison of approximate methods

Method	Does it take into account:		Does it use:
	the arrangement?	the interactions?	Maxwell–Garnett equation? (porosity)
CS	no	no	no
ERI	yes	no	yes
BPK	yes	yes	no
AD	yes	yes	no

In the next subsections, scattering cross-section values from the approximate method ($C_{method,N}$) and the exact method ($C_{Xu,N}$) will be compared. Then, the ratio K_m is defined as:

$$K_m = \frac{C_{method,N}}{C_{Xu,N}} \tag{2.12}$$

2.4.3.1 The Compact Sphere (CS) method

This method has to be mentioned because it is used as the first coarse approximation by investigators and by particle sizer manufacturers. One finds it under the name of equivalent sphere (in volume), and it will be compared with the other methods.

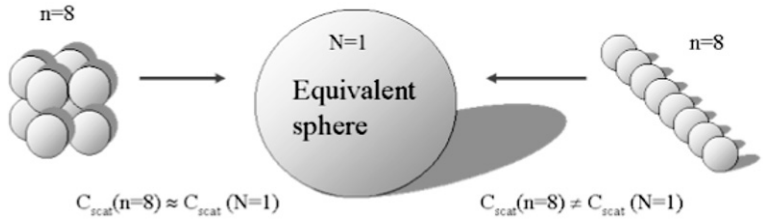


Fig. 2.10. The Compact Sphere method.

The aggregate is regarded as a full sphere, i.e. containing all the matter (Fig. 2.10). This method can be valid for aggregates of high compactness. The scattering cross-section $C_{CS,N}$ is then evaluated with the theory of Mie.

As shown in Fig. 2.11, the CS method overestimates the scattering cross-section for $x < 7$, whatever the configuration. As we will see in subsection 2.4.3.3, an aggregate can be considered as a (porous) sphere with an effective refractive index. Whatever the chosen equivalent sphere, the value of the scattering cross-section calculated by Mie theory is always smaller than the one obtained from the CS method. Moreover, this method does not take into account the interactions (interference and interaction).

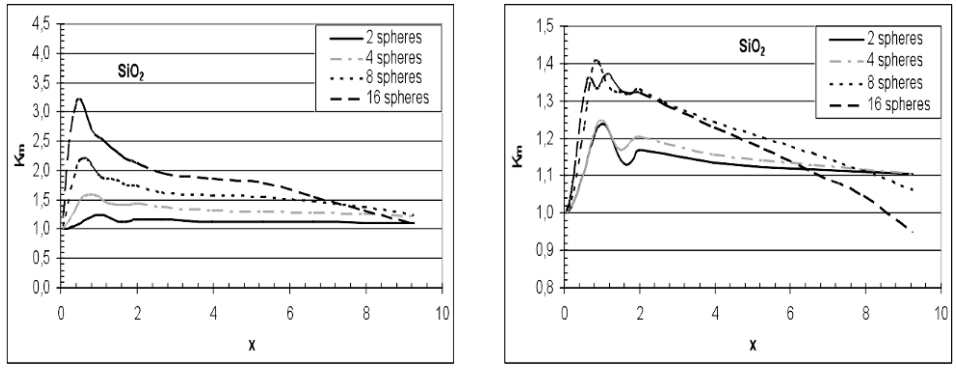


Fig. 2.11. The Compact Sphere approximation: K_m as function of the size parameter for SiO_2 (a) linear configuration, (b) compact configuration.

2.4.3.2 The Berry–Percival–Khlebtsov (BPK) method

This method originates from the work of Berry and Percival (1986) and Khlebtsov (1996). The three following points constitute the stages of the BPK method.

Step 1: Evaluation of the angular contribution of each primary particle to the scattering cross-section:

$$F_1(\theta) = \frac{C_{Mie,1}}{C_{RGD,1}} F(\theta)\,, \tag{2.13}$$

where $F(\theta)$ comes from the RGD approximation for a sphere (see Eq. (2.8b)). $F_1(\theta)$ is the corrected function for $F(\theta)$ in order to verify: $\int_0^\pi F_1(\theta)\,d\theta = C_{Mie,1}$.

Step 2: Calculation of the interferences of scattering waves for each pair of primary particles leading to a structure factor S which does not depend on polarization. The structure factor is related to the aggregate morphology through the inter-particle distances.

$$S(\theta) = \left[N + \sum_{i=j=1, i\neq j}^{N} R_{ij}(\theta) \right] \Big/ N^2\,, \tag{2.14a}$$

where

$$R_{ij}(\theta) = \frac{\sin\left(2kd_{ij}\,\sin\left(\dfrac{\theta}{2}\right)\right)}{2kd_{ij}\sin\left(\dfrac{\theta}{2}\right)} \tag{2.14b}$$

and d_{ij} is the distance between i and j particles.

Step 3: Use of a corrective coefficient d taking into account the multiple scattering (Berry and Percival, 1986)

$$d^{-1} = \left[1 - \frac{3}{2N}(-a_{i1}p_r - a_{r1}p_i)\right]^2 + \left[\frac{3}{2N}(a_{r1}p_r - a_{i1}p_i)\right]^2\,, \tag{2.15}$$

with

a_{r1} and a_{i1} are the real and imaginary parts of the first Mie coefficient a_1

$$p_r = 2\sum_{j>i,i=1}^{N} p_{r1}(kd_{ij})$$

$$p_i = 2\sum_{j>i,i=1}^{N} p_{i1}(kd_{ij})$$

$$p_{r1}(x) = \big(\cos x\, f_1(x) - \sin x\, f_2(x)\big)/x^2$$

$$p_{i1}(x) = \big(\sin x\, f_1(x) + \cos x\, f_2(x)\big)/x^2$$

$$f_1(x) = \sin x - \frac{x}{3} f(x) + \frac{1}{x}\left(f(x) - \frac{\sin x}{x}\right)$$

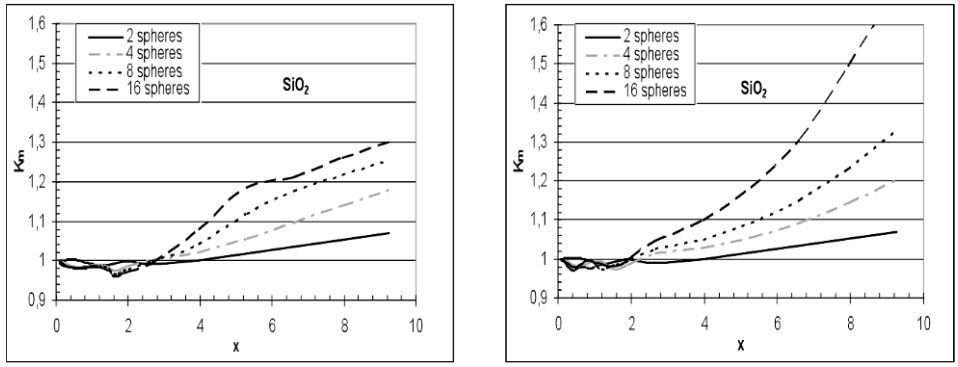


Fig. 2.12. The BPK approximation: K_m as function of the size parameter for SiO_2 (a) linear configuration, (b) compact configuration.

$$f_2(x) = \frac{\sin x}{x} - f(x)$$
$$f(x) = \frac{3}{x^3}(\sin x - x\,\cos x)$$

Conclusion of steps:

$$C_{BPK,N} = N^2 d \int_0^{\pi} F_1(\theta) S(\theta)\, d\theta \tag{2.16}$$

Fig. 2.12 shows that the BPK method is a satisfactory approximation of the exact method for the small size parameter ($x < 2$). The BPK method leads to an error smaller than 10% for a size parameter ranging between 0 and 2 in the case of SiO_2 and of Al_2O_3. For higher refractive index (i.e. TiO_2), the error increases until it reaches 30% (for more restricted size parameter range [0; 1]). The BPK method shows that the pair interactions must be taken into account only for small size parameters; their contribution in scattering cross-section calculation is less in the case of large aggregated primary particles.

2.4.3.3 The Effective Refractive Index (ERI) method

We have shown in subsection 2.4.2.2, that the location of the primary particles inside an aggregate and its shape had an effect on the scattering cross-section. The effective refractive index (ERI) method considers the shape. Knowing that the projected area of the scattering body (on the plane $(\vec{E}, \vec{H})$ of the incident wave) is a relevant parameter in optics, we consider an equivalent sphere starting from the aggregate projected area (Fig. 2.13).

Projection is carried out according to several successive planes (plane perpendicular to the incident wave vector). This corresponds to random rotation that takes place in a real situation (for instance, aggregate in a turbulent flow). Then, an average projected area $\langle S_p \rangle_0$ is calculated and an equivalent projected area

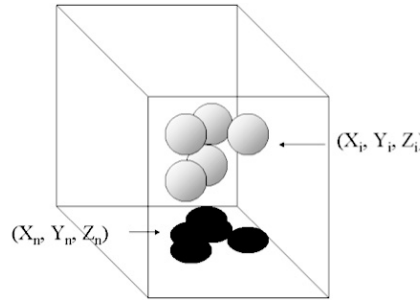


Fig. 2.13. Projected area representation.

sphere is defined. Successively, it can be deduced, the equivalent radius $a_{N,e}$, the solid volume fraction inside the sphere $\overline{\phi}_a = (Na_1^3)/a_{N,e}^3$, the effective refractive index m_a using the Maxwell–Garnett theory

$$\frac{m_a^2 - 1}{m_a^2 + 2} = \overline{\phi}_a \frac{m^2 - 1}{m^2 + 2}\,,$$

and then the extinction cross-section $C_{ERI,N}$ by means of Mie theory. This method is more efficient than the other equivalent sphere methods, because the solid volume fraction in this sphere is always high ($0.1 < \overline{\phi}_a < 1$).

ERI method behaves differently according to the configuration for small size parameter (Fig. 2.14). Indeed, Csca value calculated with this method is higher than the scattering cross-section calculated with the exact method for a linear configuration (Fig. 2.14(a)). This deviation can be taken in consideration and calculation has to be corrected in order to reduce the deviation between ERI and exact methods.

Jacquier and Gruy (2007a, b) proposed a corrective factor $F(x, d_1/a)$ for the scattering cross-section $C_{ERI,N}$. This is written as:

$$C_{ERI,N}^{corr} = C_{ERI,N}/F(x, d_1/a)\,, \tag{2.17}$$

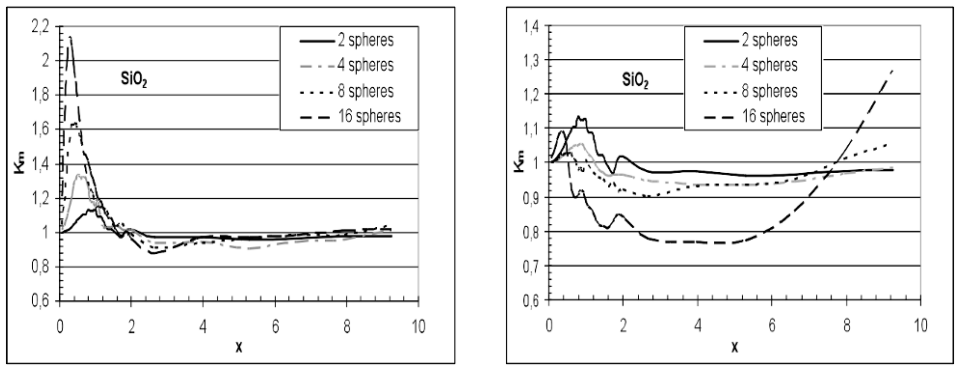


Fig. 2.14. The ERI approximation: K_m as function of the size parameter for SiO_2 (a) linear configuration, (b) compact configuration.

where d_1 is a morphological parameter defined as:

$$d_1 = \frac{1}{N(N-1)} \sum_{i,j} d_{ij} \tag{2.18}$$

2.4.3.4 The Anomalous Diffraction (AD) method

In subsection 2.4.2.2 it has been already mentioned that morphology becomes more important for large size parameters. The anomalous diffraction approximation, clarified in subsection 2.3.3 (Van de Hulst, 1981), accounts for the aggregate morphology by means of the intercept (chord) of a light ray and the aggregate (Fig. 2.15).

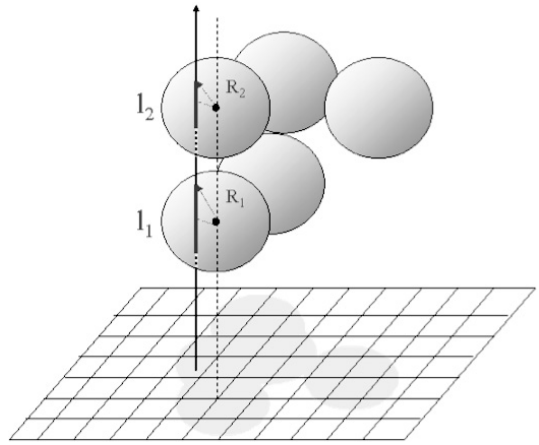


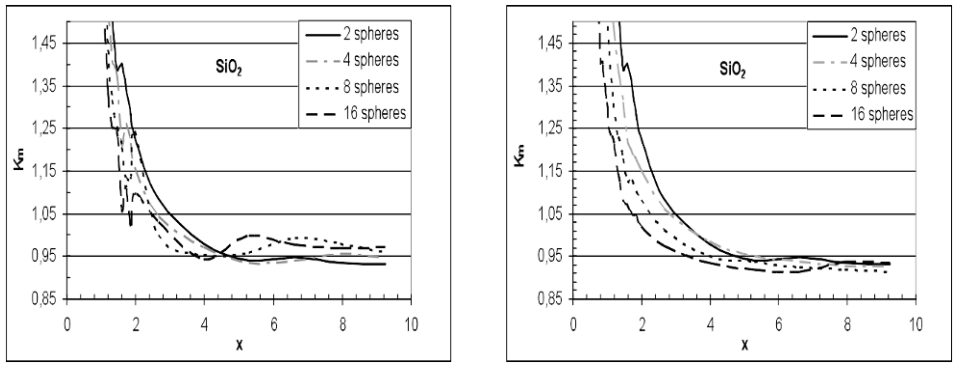
Fig. 2.15. Definition of a chord.

The various possible chords $l_1, l_2 \ldots$ were evaluated and introduced as $l = \sum_i l_i$ into Eq. (2.19)

$$C_{AD,N} = 2 \iint_{[S_p]} \left(1 - \cos \frac{2\pi}{\lambda} l(m-1)\right) dy\, dz = 2 \iint_{[S_p]} \left(1 - \cos x(l/a)(m-1)\right) dy\, dz\,. \tag{2.19}$$

$[S_p]$ is the projection plane. Details or examples on expressions relating l/a and (y, z) can be found in Yang et al. (2004) and Gruy and Jacquier (2008).

This calculation is repeated while rotating the aggregate or changing the projection plane. So, a mean value of scattering cross-section is deduced (Fig. 2.16). As expected, the AD method is not suitable for small size parameters but proves to be a good approximation for a size parameter higher than 2. It is important to recall that AD is strongly related to the configuration (morphology) since it includes in its formulation the chord length distribution of the aggregate.



a b

Fig. 2.16. The AD approximation: K_m as function of the size parameter for SiO_2 (a) linear configuration, (b) compact configuration.

2.4.3.5 Summary

Approximate methods facilitate the estimation of the aggregate scattering cross-section in a short computation time. Jacquier and Gruy (2007a) evaluated the performance of these four approximate methods with respect to the exact method:

- Methods replacing an aggregate by a compact sphere (CS) are inappropriate.
- The BPK (Berry–Percival–Khlebtsov) method is valid for $0 < x < 2$ with an error which increases with the material refractive index.
- The corrected ERI (Effective Refractive Index) method is the approximate method capable of efficiency over all size parameters. The error for a scattering cross-section is always smaller than 5%.
- The AD (Anomalous Diffraction) method works fairly well for $2 < x < 10$ and is less sensitive to refractive index variation.

2.4.4 Application: turbidity versus time during the agglomeration process

As mentioned in section 2.2, nucleation and growth lead to (primary) particles with a size between 0.1 μm and 10 μm. Then, these particles collide and agglomerate by Brownian motion and/or local shear. So, let us consider agglomeration of small monosized primary particles in a homogeneous suspension. Agglomeration proceeds as a bimolecular reaction, the kinetic constant of which can be expressed in terms of known quantities. Generally, the kinetic constant is a function of the sizes of the two colliding particles. However, in the case of Brownian agglomeration or shear agglomeration (but not for shear aggregation, i.e without consolidation of the particle cluster), the kinetic constant K_{ag} weakly depends on the particle size, so that we may consider it as not dependent on particle size. Following Kruyt (1952), modeling of agglomeration with constant K_{ag} leads to simple expressions for number concentrations in the agglomerate:

$$N_j = N_0(t/t_c)^{j-1}(1+t/t_c)^{-j-1}\,. \tag{2.20a}$$

N_j is the number concentration in agglomerate consisting of j primary particles. N_0 is the number concentration in primary particles at time $t = 0$. There is no agglomerate at $t = 0$. t_c is the characteristic time of agglomeration. It obeys:

$$t_c = 2/(K_{ag}N_0)\,. \tag{2.20b}$$

For instance, $t_c = 3\mu/(4kTN_0\alpha)$ for Brownian agglomeration. T, k and μ are the temperature, the Boltzmann constant and the dynamic viscosity, respectively. α is a non-dimensional parameter representing the agglomeration efficiency ($0 < \alpha < 1$). For the sake of simplicity, we consider this parameter as a constant throughout the agglomeration process.

It will be pointed out that t_c and then N_j do not depend on the agglomerate morphology. The previous expressions are approximate, but are considered as a first and realistic approach to agglomeration dynamics.

At a given time, the turbidity of the suspension contains the contribution of each j-agglomerate:

$$\tau(\lambda,t) = \sum_{j=0}^{\infty} N_j(p_1^j \ldots p_p^j, t)C_{ext}(\lambda, m, p_1^j \ldots p_p^j)\,. \tag{2.21}$$

Following the ERI method, the internal coordinates relevant for scattering cross-section are a (the primary particle radius), j, $\langle S_p\rangle_0$. Even if the characteristic time does not depend on the morphology, it appears that large agglomerates have a fractal-like structure. Depending on the agglomeration mechanism, simulations give values of fractal dimension between 1.8 and 2.6. Due to restructuring of agglomerates, the fractal dimension is larger than 2. As the fractal dimension is larger than 2, the outer radius of the agglomerate is equal to the radius $a_{S,j}$ of the 'projected area' equivalent sphere. Small agglomerates do not have the fractal-like structure. However, we have shown (Gruy, 2001) that they can be described by means of a power law relating $a_{S,j}$ and j:

$$\frac{a_{S,j}}{a} = \left(\frac{\langle S_p\rangle_0}{\pi a^2}\right)^{1/2} = \left(\frac{j}{S_r}\right)^{1/D_F}\,. \tag{2.22}$$

Later on, we will consider Eq. (2.22) suitable for a wide range of primary particle numbers. Then,

$$\tau(\lambda,t) = \sum_{j=0}^{\infty} N_j(t,t_c)C_{ext}(\lambda, m, j, a, D_F)\,. \tag{2.23}$$

Figs 2.17 and 2.18 represent the change of turbidity (normalized by $\tau_0(\lambda) = N_0C_{ext}(\lambda, m, a)$) with time (normalized by t_c) for agglomeration of silica ($m = 1.08$). Figs 2.17 and 2.18 show the effect of two fractal dimensions ($D_F = 2.1$; $D_F = 2.5$) and two primary particle radii ($a = 0.1\,\mu$m; $a = 1\,\mu$m) respectively at $\lambda = 0.4\,\mu$m and $\lambda = 0.8\,\mu$m.

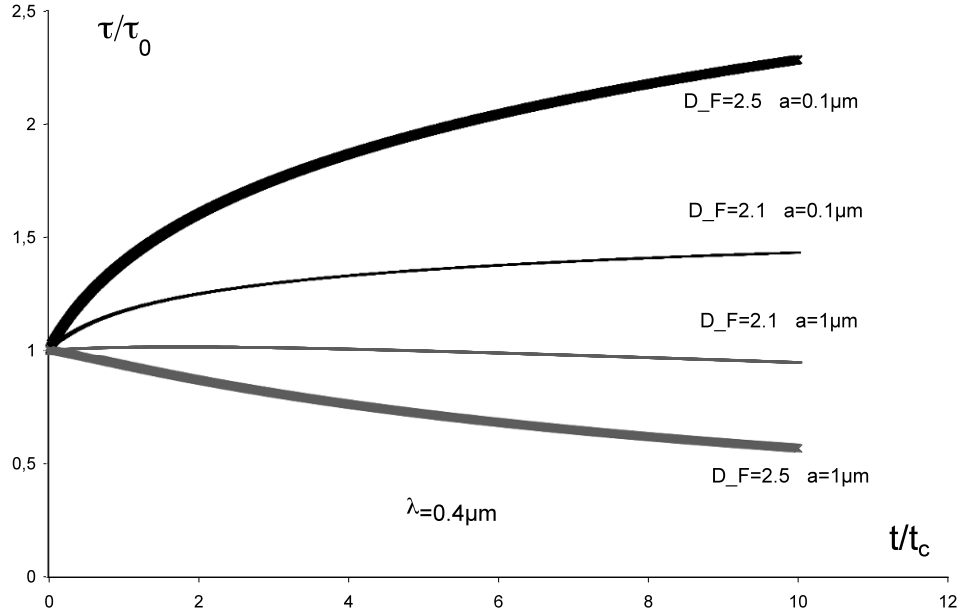


Fig. 2.17. Normalized turbidity versus time; agglomeration of silica in water; $\lambda = 0.4\,\mu$m.

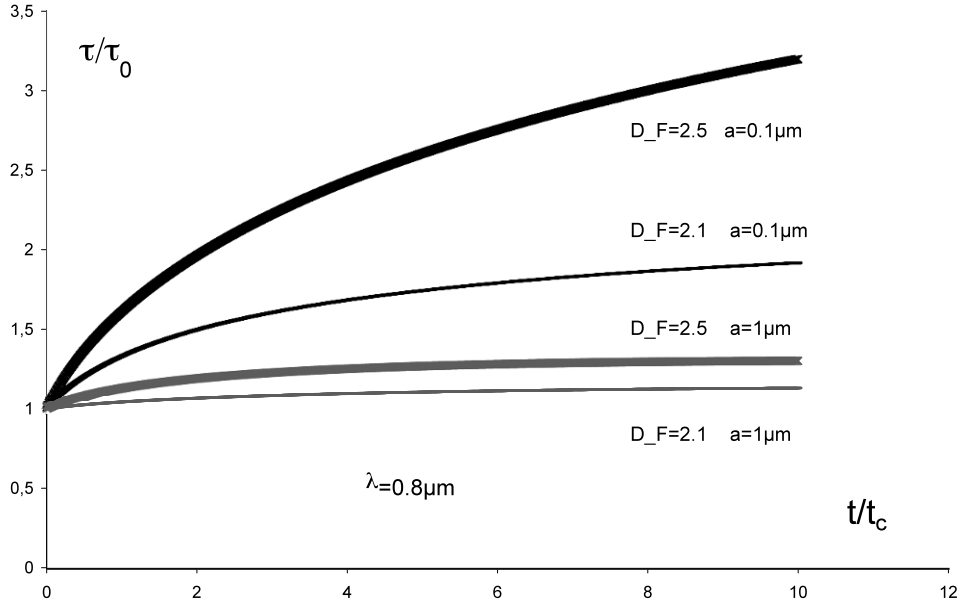


Fig. 2.18. Normalized turbidity versus time; agglomeration of silica in water; $\lambda = 0.8\,\mu$m.

2.5 Approximation for radiation pressure cross-section

2.5.1 Introduction

Often, precipitation leads to concentrated suspensions. On the one hand, transmitted light intensity becomes very weak making backscattering sensors more suitable than turbidimetric ones. On the other hand, multiple scattering takes place. Whatever the considered signal (backscattered, side-scattered or transmitted light), the interpretation has to account for multiple scattering. The most popular theory which considers this phenomenon is the radiative transfer theory, particularly its diffusion approximation (Ishimaru, 1978). The relevant phenomena associated with backscattering measurement are either coherent or incoherent ones. The first type result from interference caused by the double passage of the wave through the same particle (Tsang and Ishimaru, 1984, 1985; Wolf et al., 1988; Akkermans et al., 1988; de Wolf, 1991; Helfenstein et al., 1997). The angular width of the measured intensity peak is proportional to the transport mean free path $L_{tr} = (NC_{pr})^{-1}$, where N is the particle number concentration and C_{pr} is the radiation pressure cross-section. The second type only considers the multiple scattering: scattered light intensity is also a function of the transport mean free path. Theoretical calculations were compared to on- (off-) line experimental data for transmittance (Ishimaru et al., 1983) and retroreflectance (Kuga and Ishimaru, 1984; Nichols et al., 1997) experiments with suspensions of spherical beads, and a fairly good agreement between both was found.

The radiation pressure cross-section is expressed as a function of the extinction and scattering cross-sections, and also the asymmetry parameter ($g = \langle \cos\theta \rangle$):

$$C_{pr} = C_{ext} - C_{sca} \cdot g \,.$$

In the case of non absorbing material:

$$C_{pr} = C_{sca}(1 - g) \,.$$

The asymmetry parameter for spheres has been calculated and analytical or empirical expressions have been derived in the framework of different approximations, e.g. geometrical optics and RGD approximations. Empirical expressions for large randomly oriented non-spherical particles were also proposed (see Kokhanovsky, 2001). Rigorous calculations were achieved by different authors (see, for instance, Xu (1998b)).

Up to now, calculations of radiation pressure cross-section for aggregates have been mainly motivated by the calculation of forces acting on interstellar dust illuminated by stars (Kimura and Mann, 1998; Kimura et al., 2002; Iati et al., 2004). Radiation pressure plays a key role in the dynamical behavior of submicrometer-size grains in the stellar radiation and gravitational fields. Kimura and Mann (1998) studied aggregates composed of 256 primary particles, the radius of which is 0.01 μm and that are illuminated by visible light and infrared. The considered materials are silicate and amorphous carbon as representatives of weakly and strongly absorbing materials, respectively. Calculations for randomly oriented fractal-like aggregates (with $D_F = 2$ and $D_F = 3$) were performed by means of the DDA method. Authors showed that the asymmetry parameter smoothly increases with increasing

size parameter x (decreasing wavelength) of the primary particle and increases as the fractal dimension decreases if $x < 0.16$ (for $x > 0.16$, $g \simeq 0.7$). The asymmetry parameter for aggregates is higher than for volume-equivalent spheres, irrespective of the constituent material. The authors point out that aggregates with small fractal dimension present a large fluctuation in g for different aggregate orientations. Kimura et al. (2002) extended the previous study to larger aggregates ($N < 2048$). They compare radiation pressure cross-sections calculated from the DDA method, Mie theory applied to volume-equivalent sphere (CS method) and Mie theory combined with the Bruggeman mixing rule. CS is a rough approximation for the two materials and two fractal dimensions. Mie/Bruggeman approximation is a good approximation for compact aggregates, but performs weakly for loose aggregates, especially with non-absorbing primary particles. The authors showed that C_{pr} is less dependent on the porosity of aggregates while the values strongly vary with the material composition. Iati et al. (2004) computed, through the T-matrix method, optical properties of cosmic dust grains. Grains are aggregates consisting of 31 non-identical spheres. Materials are also silicate and amorphous carbon. Primary particle size distribution is assumed to be Gaussian-like. The radius of the volume-equivalent sphere is equal to $0.14\,\mu$m. For both materials, aggregation leads to a sharp increase in the extinction and radiation pressure cross-sections. Subsections 2.5.2 and 2.5.3 are respectively devoted to the main features of C_{pr} for aggregates and approximate methods for estimating C_{pr}.

2.5.2 Main features of radiation pressure cross-section

2.5.2.1 Single sphere

The variation of the asymmetry parameter is presented for spherical particles of various chemical compositions in Fig. 2.19.

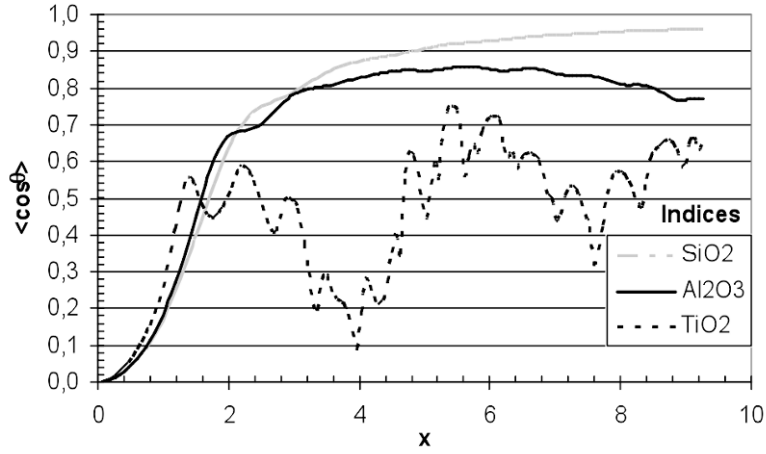


Fig. 2.19. Asymmetry parameter for three materials as function of the size parameter.

The asymmetry parameter is close to zero for very small size parameters, whatever the refractive index. Then, it increases to a plateau close to 1 in the case of SiO_2 and Al_2O_3, i.e. the light is scattered predominantly in the forward direction. A spherical particle of TiO_2 has a mean scattering angle which varies less monotonously according to the size parameter. Indeed, for a size parameter equal to 4, the asymmetry parameter is close to zero, the scattering can be then described as dipole-like, while for a size parameter of about 6 the scattering seems to happen in a privileged direction. Fig. 2.20 represents C_{pr} as a function of the size parameter for a sphere and the three different materials. In the size parameter range [0; 10], the radiation pressure cross-section increases as the optical refractive index increases.

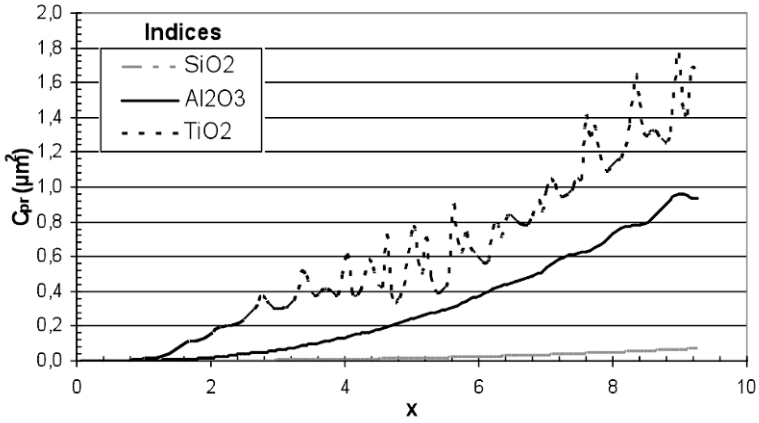


Fig. 2.20. Radiation pressure cross-section for three materials as function of the size parameter.

2.5.2.2 Aggregate of spheres

The variations of C_{pr} for an aggregate have been examined according to: the number of primary particles, their size parameter, the relative optical refractive index and the aggregate shape. Simulations were performed by means of the GMM method (Xu, 1998b).

The simulation results are presented as previously: effect of the number of primary particles within the aggregate and effect of the aggregate shape on the function $P_{Xu}(x)$. P_{Xu} is the ratio between C_{pr} value of an aggregate ($C_{pr,Xu,N}$) and one of its N primary particles considered as independent $NC_{pr,Mie,1}$:

$$P_{Xu} = \frac{C_{pr,Xu,N}}{NC_{pr,Mie,1}} .$$

2.5.2.2.1 Effect of the number of primary particles on the radiation pressure cross-section

Fig. 2.21 represents P_{Xu} as a function of the size parameter for a chain-aggregate of SiO_2 primary particles. The variation of C_{pr} with primary particle size parameter is similar to that one corresponding to the scattering cross-section. Two size parameter ranges can be defined. In the x-range [0; 2], constructive interferences and multiple scattering (or interaction between primary particles) are important. In the x-range [2; 10], the radiation pressure cross-section is close to the C_{pr} of a set of spheres without interaction. However, multiple scattering still occurs at some extend.

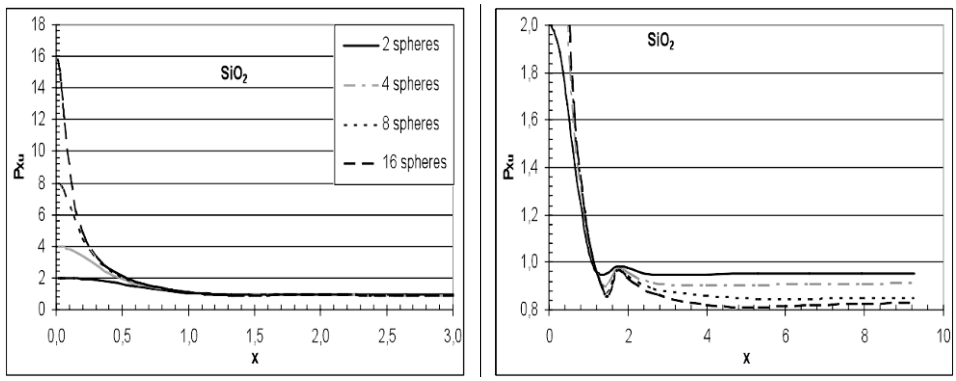


Fig. 2.21. P_{Xu} as a function of the size parameter for the linear configuration with N primary particles (2, 4, 8, 16) and for SiO_2: (a) for the range $x \in [0, 3]$ and (b) $x \in [0, 10]$.

However, the radiation pressure cross-section of an aggregate made up of primary particles whose size parameter is higher than 1.5 seems to be proportional (by a factor $p(N, x)N$) to the primary particle radiation pressure cross-section. $p(N, x)N$ is a weak function of x for SiO_2 (more dependent on x for TiO_2). As expected, $P_{Xu}(x)$ is similar to $K_{Xu}(x)$ in the x-range [0; 1] because the asymmetry factor of the primary particle is smaller than 0.25. But, $P_{Xu}(x)$ must not be related to $K_{Xu}(x)$ in the x-range [1; 10].

2.5.2.2.2 Effect of the aggregate morphology on radiation pressure cross-section

Fig. 2.22 represents P_{Xu} for different configurations of aggregates consisting of four primary particles arranged according to Fig. 2.9. Similar variations are obtained. The deviation between the two extreme configurations is about 10.7%, which is a smaller value than that obtained with C_{sca}. But C_{pr} is a little more sensitive to configurations which are close each other, since the average deviation between the p1 and p2 configurations is about 1.3% compared with 0.8% for C_{sca}.

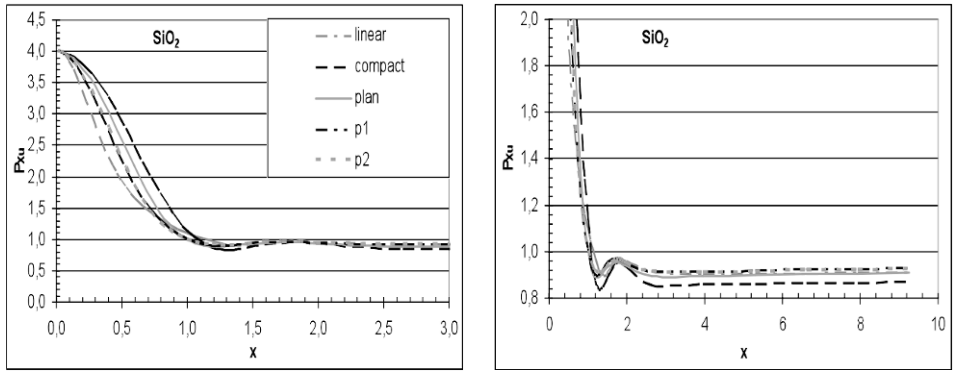


Fig. 2.22. P_{Xu} as a function of the size parameter for different configurations of 4-aggregates (linear, compact, plane, p1, p2) for SiO_2.

2.5.3 Approximate methods for aggregates

2.5.3.1 The ERI method

In the same way that we showed the effectiveness of the ERI approximate method, for calculation of C_{sca}, we evaluated (Jacquier and Gruy, 2007b) its performances for the calculation of C_{pr}. The ratio of the radiation pressure cross-sections obtained on the one hand with the exact (GMM) method and on the other hand with the ERI method is denoted L_m:

$$L_m = \frac{C_{pr,ERI,N}}{C_{pr,Xu,N}} .$$

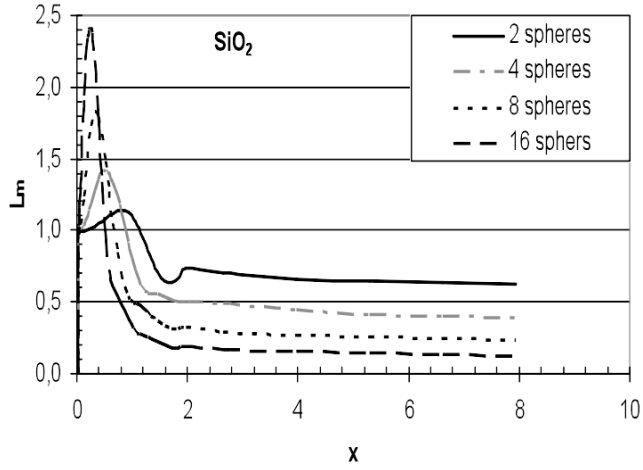


Fig. 2.23. L_m function with $x \in [0; 10]$ for linear aggregate (SiO_2).

The deviation of L_m from 1 (Fig. 2.23) leads to the search for a corrective factor in order to reduce it. Jacquier and Gruy (2007b) proposed a corrective factor as a multi-parameter function $G(x, N, d_1/a)$ for the radiation pressure cross-section $C_{pr,ERI,N}$. Thus, the corrected radiation pressure cross-section obeys the expression:

$$C^{corr}_{pr,ERI,N} = C_{pr,ERI,N}/G(x, N, d_1/a)\,. \tag{2.24}$$

Later on, this method is called ERI/G.

2.5.3.2 Other methods

We noticed in subsection 2.5.2.2 that P_{Xu} of any configuration of soft particles does not vary with x for x higher than 1.5. The value of $P_{Xu(x>1.5)}$ depends on the aggregate morphology that can be characterized through N and d_1/a. However, $P_{Xu(x>1.5)}$ may be a weak function of x for hard material (Fig. 2.24). We can observe that variations of $P_{Xu(x>1.5)}$ are similar to those for a two-sphere aggregate.

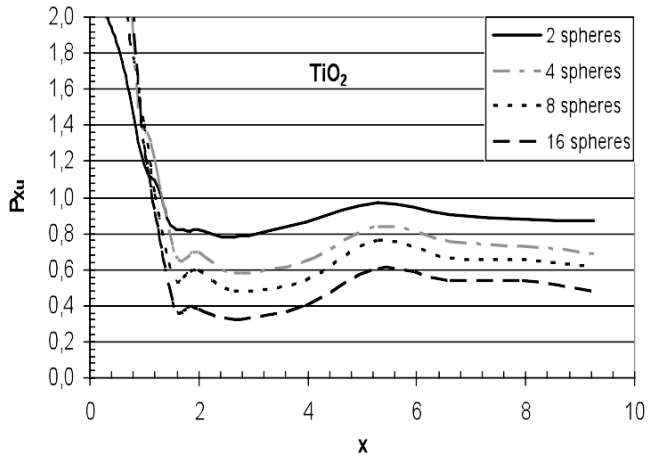


Fig. 2.24. P_{Xu} as a function of the size parameter for the compact configuration with N primary particles (2, 4, 8, 16) and for TiO_2: $x \in [0, 10]$.

These comments suggest the two approximate methods:

- method Pp1: $C_{pr,Xu,N}$ proportional to C_{pr} of a primary particle: $C_{pr,Mie,1}$
- method Pp2: $C_{pr,Xu,N}$ proportional to C_{pr} of a doublet: $C_{pr\,Xu,2}$

The proportionality factors, denoted respectively p_1 and p_2, can be expressed as a function of a single parameter: $\beta = N/\sqrt{d_1/a}$. Corresponding expressions can be found in (Jacquier and Gruy, 2007b). Table 2.3 presents the performances of ERI/G, Pp1 and Pp2. It appears that the ERI/G method is not as efficient as Pp1 and Pp2 but ERI/G presents the biggest advantage to be used over all the size range.

Table 2.3. Approximate method performance for the calculation of C_{pr} for different materials with $x \in [1.5; 10]$. m and σ are respectively the mean value and the standard deviation for the corresponding dataset

method	Linear configuration						Compact configuration					
material	SiO_2		Al_2O_3		TiO_2		SiO_2		Al_2O_3		TiO_2	
	m	σ	m	σ	m	σ	m	σ	m	σ	m	σ
Pp1	1.01	0.05	1.05	0.11	1.06	0.17	1.01	0.04	1.00	0.11	1.08	0.17
Pp2	0.98	0.04	0.96	0.08	0.94	0.12	0.97	0.04	0.93	0.10	0.93	0.11
ERI/G	0.93	0.17	0.90	0.15	1.07	0.23	0.92	0.13	0.90	0.11	1.06	0.30

2.5.4 Conclusion

We have presented different ways to calculate approximately the radiation pressure cross-section of aggregates. The corresponding expressions can be used to study dense suspensions. For instance, Tontrup et al. (2000) performed an experimental work about the aggregation of TiO_2 micro-particles in water by using a backscattering sensor. They deduced the change of the transport mean free path with time. SEM observations showed that the aggregates contain few primary particles. Approximations could be used to determine some characteristics of the aggregates.

2.6 Scattering properties versus geometrical parameters of aggregates

The main question that appears when studying the formation of particles or aggregates is: which is the relevant morphological parameter related to the measurement? The answer mainly depends on the particle size and is partially included in theories and modeling leading to scattering cross-section calculations.

So, when we consider the Mie theory for homogeneous spheres, the solving method and the results are dependent only on the relative refractive index and the boundary conditions for the Helmholtz equation. From a geometrical point of view, the mathematical function describing the particle surface is the relevant parameter. The case of non-spherical convex bodies is similar. As the physics is always based on the Maxwell and Helmholtz equations, the corresponding solution for a natural incident light depends only on the body surface that is characterized by the equation $f(x, y, z) = 0$.

If we are interested in the orientation average of the scattering cross-section, a function describing the body and being invariant to rotation will be preferred. So, the pair distance distribution density (PDDD) could be an interesting approach to describing the shape. It is a well-known function in physics and can be defined for liquids as:

$$dN = g(r)4\pi r^2\, dr\,.$$

dN represents the number of molecules distant from a given (tagged) molecule with the distance in the range $[r, r + dr]$. In the case of liquid, the distribution is nearly

isotropic. This function clearly appears in RGD approximation for convex bodies (distribution density is connected to $|f|^2$) and BPK approximation for multi-sphere aggregate (in Eqs (2.14a) and (2.14b)).

In the first case (RGD), we consider any pair of volume elements in the scattering body. The pair distance distribution density is a continuous function of the distance between volume elements. In our context, we chose the notation $D_p(r)$. Then, the orientation-averaged scattering cross-section can be written as (Gruy, 2009):

$$C_{RGD,1} = \frac{2}{3\pi} k^4 V^2 (m-1)^2 \int_{R_{min}}^{R_{max}} F(kr) D_p(r)\, dr \tag{2.25}$$

with

$$F(x) = 3\big[\cos(2x)(-1+5x^{-2}-3x^{-4})+\sin(2x)(2x^{-1}-6x^{-3})+1+x^{-2}+3x^{-4}\big]/(4x^2)\,.$$

The distribution density function is normalized:

$$\int_{R_{min}}^{R_{max}} D_p(r)\, dr = 1 \tag{2.26}$$

Fig. 2.25 presents the function $D_p(r)$ for a sphere and various spheroids. The pair distribution function for a sphere with radius a obeys the expression:

$$aD_p(r) = D_{p,u}(u) = \frac{3}{16} u^2 (u^3 - 12u + 16) \tag{2.27}$$

with $u = r/a$ and $0 < u < 2$.

In the second case (BPK), equations contain the inter-particle distance d_{ij}. This function is not continuous; as far as a cluster of point scatterers is concerned:

$$D_p(r) = \frac{1}{N(N-1)} \sum_{i,j} \delta(r - d_{ij})\,. \tag{2.28}$$

δ is the Dirac function.

A particular case is the fractal-like aggregate, the PDDD of which obeys the equation (continuous form):

$$D_p(r) \propto r^{D_F - 3}\,. \tag{2.29}$$

According to subsection 2.4.3.2, the BPK approximation gives good results when the size parameter of the primary particle is smaller than 2. Thus, the PDDD is the relevant morphological parameter.

It has been shown by Gruy (2009) that this function associated with BP approximation (Berry and Percival, 1986) for aggregates of Rayleigh scatterers allows for an estimation of the scattering cross-section of convex bodies. This method is proved efficient for spheres and spheroids as the scattering efficiency is smaller than 1 and as the material is optically either soft or hard.

The non-corrected ERI method is based only on the average projected area $\langle S_p \rangle_0$ of the body. However, it is not suitable for elongated aggregates with small size parameter. $\langle S_p \rangle_0$ can be expressed as a function of the number of primary particles and of the mean inter-particle distance (Jacquier and Gruy, 2008a,b):

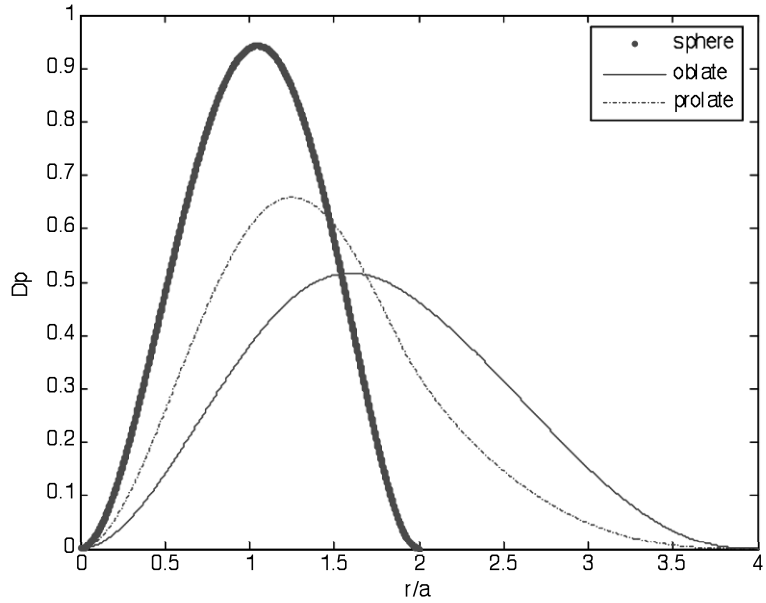


Fig. 2.25. Pair distance distribution density for sphere, oblate spheroid (axis ratio equal to 2) and prolate spheroid (axis ratio equal to 2). a is the smallest semi-axis length.

$$\langle S_p \rangle_0 = \pi R_p^2 \,, \tag{2.30}$$

with $R_p = a(d_1/2a)^{1/5} N^{1/3} d_E^{1/8}$ and $N < 100$; d_E is the space dimension; d_1 is the relevant morphological parameter. It is directly related to the first moment of the distribution $D_p(r)$:

$$\begin{aligned} M_1 &= \int_0^\infty D_p(r) r \, dr \\ &= \int_0^\infty \frac{1}{N(N-1)} \sum_{i,j} \delta(r - d_{ij}) r \, dr \\ &= \frac{1}{N(N-1)} \sum_{i,j} d_{ij} = d_1 \,. \end{aligned} \tag{2.31}$$

It would be possible to choose other moments of the distribution for describing geometrical and optical properties of aggregates. For instance, the second-order moment is directly related to the gyration radius, that is a well-known parameter used to define a fractal-like aggregate. However, there was no noticeable change and thus no improvement was found when choosing another mean distance definition for the aggregate. Thus, we chose the lowest-order distribution moment. The corrected ERI method also uses the d_1 distance parameter.

For large size parameter ($x > 2$), the AD approximation becomes efficient. In this case, the relevant line is the chord. Expression of the average scattering cross-section can be rewritten by introducing the chord length distribution $D_l(l)$ (Jacquier and Gruy, 2008a,b):

$$\langle C_{AD,N} \rangle = \left\langle 2 \iint\limits_{[S_p]} \big(1 - \cos kl(m-1)\big)\, dy\, dz \right\rangle$$
$$\simeq 2\langle S_p \rangle \int_0^{l_{max}} \big(1 - \cos kl(m-1)\big) D_l(l)\, dl \tag{2.32}$$

The chord length distribution (CLD) is defined as follows: $D_l(l)\, dl$ represents the number fraction of the chord length in the range $[l, l + dl]$. Thus, $D_l(l)$ obeys the normalization equation:

$$\int_0^{l_{max}} D_l(l)\, dl = 1\,. \tag{2.33}$$

l_{max} is the maximum chord length of an aggregate.

Fig. 2.26 presents the $D_l(l)$ function for an aggregate consisting of 16 primary particles. One observes three very distinct peaks or modes, each one characterized by a chord length range:

- $[0; 2a]$ corresponds to primary particles (distribution density $D_{l,1}(l)$)
- $[2a; 4a]$ corresponds to pair of particles in contact (distribution density $D_{l,2}(l)$)
- $[4a; l_{max}]$ corresponds to the aggregate superstructure (distribution density $D_{l,G}(l)$)

$D_l(l)$ contains the contributions of each distribution density. These are weighted by the coefficients ξ_1, ξ_2, ξ_3:

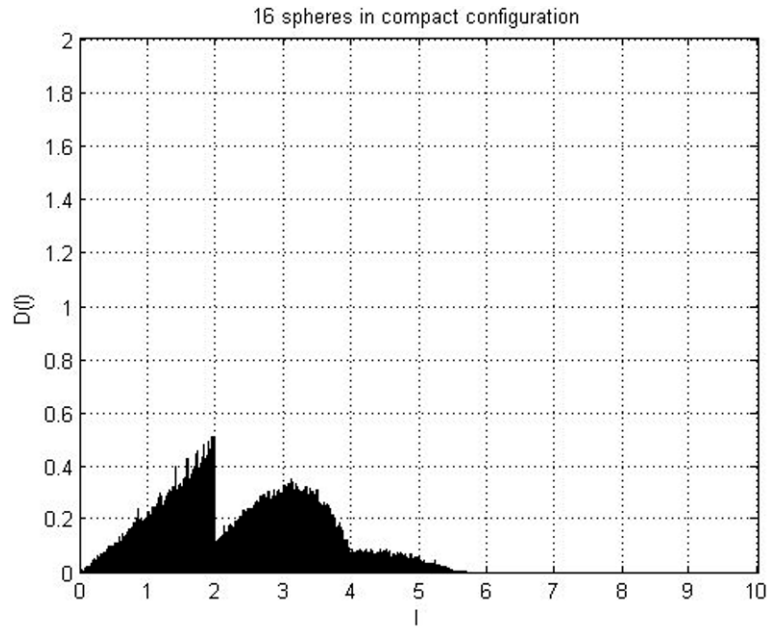


Fig. 2.26. $D_l(l)$ for compact aggregate with 16 primary particles. l is normalized by the radius of the primary particle.

$$D_l(l) = \xi_1 D_{l,1}(l) + \xi_2 \frac{D_{l,2}(l \in [2a; 4a])}{\int_{2a}^{4a} D_{l,2}(l)\, dl} + \xi_G \frac{D_{l,G}(l \in [4a; l_{max}])}{\int_{4a}^{l_{max}} D_{l,G}(l)\, dl} \tag{2.34a}$$

with

$$\xi_1 + \xi_2 + \xi_G = 1\,. \tag{2.34b}$$

The distribution densities $D_{l,1}(l)$ and $D_{l,2}(l)$ are given by analytical expressions (Jacquier and Gruy, 2008a,b). $D_{l,G}(l)$ is an empirical function, the same for all the aggregates. Only weighting coefficients depend on the aggregate morphology. Results shown in Fig. 2.26 for a particular aggregate are representative for aggregates with primary particle number up to one hundred. Coefficients ξ_1, ξ_2, ξ_G depend on N, d_1, d_E (Jacquier and Gruy, 2008a,b).

The calculation of the scattering cross-section using Eqs (2.32) to (2.34b)) is much faster than that based on Eq. (2.19).

Certain presented approximations are characterized by a decoupling or separation between optics and geometry. This separation allows for a faster calculation of the optical properties. To our knowledge, the relationship between chord length distribution (as defined by Fig. 2.15) and pair distance distribution is not trivial, particularly for aggregates, and requires complementary works in the field of integral geometry. Moreover, the transition between the different geometrical characteristics, i.e PDD and CLD, as the primary particle size increases is not yet quantitatively understood.

2.7 Conclusion

The analysis of turbidimetric data during the precipitation process is a challenge for researchers working in the field of light scattering by particles. The variety of sizes, shapes and optical contrast requires several approaches for the calculation of their optical properties. Performance criteria are the calculation speed and the accuracy fitted with the measurement accuracy. Accurate calculations performed with sophisticated numerical methods will always be needed and used for the purpose of validation. Difficulties remain for certain particles with a complex morphology. For instance, one observes precipitated zinc sulphide particles in the size range [0.5 μm; 5 μm] exhibiting sand rose (i.e., gypsum flower) morphology. The typical multi-scale morphology of many precipitated particles firstly needs tools coming from integral geometry in order to be described with a minimum number of parameters. Knowing this parameter set, optical properties will be calculated with exact theories. The parameter number coming from geometrical analysis can be reduced when the optical properties of the particles are appropriately considered for the formulation of approximate theories to the calculation of scattering properties. An example for such an approach has been presented, but further advances are needed.

Acknowledgment

The authors thank Frank Babick from the Technical University of Dresden (Germany) for the helpful discussions and the final revision of this text.

References

Akkermans, E., P. E. Wolf, R. Maynard, G. Maret, 1988: Theoretical study of the coherent backscattering of light by disordered media, *J. Phys. Fr.*, **49**, 77–98.

Asano, S., 1979: Light scattering properties of spheroidal particles, *Appl. Optics*, **18**(5), 712–723.

Asano, S. and M. Sato, 1980: Light scattering by randomly oriented spheroidal particles. *Appl. Optics*, **19**(6), 962–974.

Asano, S. and G. Yamamoto, 1975: Light scattering by spheroidal particle, *Appl. Optics*, **14**(1), 29–49.

Auger, J.-C., R. G. Barrera, B. Stout, 2003: Scattering efficiency of clusters composed by aggregated spheres, *J. Quant. Spectr. Rad. Transfer*, **79–80**, 521–531.

Auger, J.-C., B. Stout, V. Martinez, 2005: Scattering efficiency of aggregated clusters of spheres: dependence on configuration and composition, *J. Opt. Soc. Am.*, **22**(12), 2700–2708.

Berry, M. V. and I. C. Percival, 1986: Optics of fractal clusters such as smoke, *Opt. Acta*, **33**(5), 577–591.

Bohren, C. F. and D. R. Huffman, 1998: *Absorption and Scattering of Light by Small Particles*, Wiley-VCH, Berlin.

Cameirao, A., R. David, F. Espitalier, F. Gruy, 2008: Effect of precipitation conditions on the morphology of strontium molybdate agglomerates, *J. Cryst. Growth*, **310**, 4152–4162.

Chylek, P. and J. D. Klett, 1991: Absorption and scattering of electromagnetic radiation by prismatic columns: anomalous diffraction approximation, *J. Opt. Soc. Am.*, **8**, 274–281.

Coccioli, R., T. Itoh, G. Pelosi, P. P. Silvester, 1996: Finite elements methods in microwaves: a selected bibliography, *Antennas Propag. Mag.*, **38**, 34–48.

Coudun, C., E. Amblard, J.-C. Guihaumé and J.-F. Hochepied, 2007: Nanostructured particles by controlled precipitation techniques; example of nickel and cobalt hydroxides, *Catal. Today*, **124**, 49–54.

de Wolf, D. A.,1991: Backscatter enhancement: random continuum and particles, *J. Opt. Soc. Am. A*, **8**, 465–471.

Draine, B. T. and P. J. Flatau, 1994: Discrete-dipole approximation for scattering calculations, *J. Opt. Soc. Am.*, **11**(4), 1491–1499.

Elimelech, M., J. Gregory, X. Jia, R. Williams, 1995: *Particle deposition and aggregation, measurement, modelling and simulation*, Butterworth-Heinemann Ltd, Oxford.

Farafonov, V. G., V. B. Il'in, T. Henning, 1999: A new solution of the light scattering problem for axisymmetric particles, *J. Quant. Spectr. Rad. Transfer*, **63**, 205–215.

Fuller, K. A. and G. W. Kattawar, 1988a: Consummate solution to the problem of classical electromagnetic scattering by an ensemble of spheres. I: Linear chains, *Opt. Lett.*, **13**(2), 90–92.

Fuller, K. A. and G. W. Kattawar, 1988b: Cosummate solution to the problem of classical electromagnetic scattering by an ensemble of spheres II: Clusters of arbitrary configuration, *Opt. Lett.*, **13**(12), 1063–1065.

Gruy, F., 2001: Formation of small silica aggregates by turbulent aggregation, *J. Colloid Interf. Sci.*, **237**, 28–39.

Gruy, F., 2009: Light scattering cross-section as a function of pair distribution density, *J. Quant. Spectr. Rad. Transfer*, **110**, 240–246.

Gruy, F. and S. Jacquier, 2008: The chord length distribution of a two-sphere aggregate, *Comp. Mater. Sci.*, **44**, 218–223.

Harrington, R. F., 1968: *Field computation by moment methods*, New York, Macmillan.

Helfenstein, P., J. Ververka, J. Hillier, 1997: The lunar opposition effect, a test of alternative models, *Icarus*, **128**, 2–14.

Hovenier, J. W., K. Lumme, M. I. Mishchenko, N. V. Voshchinnikov, D. W. Mackowski, J. Rahola, 1996: Computations of scattering matrices of four types of non-sphericles using diverse methods, *J. Quant. Spectr. Rad. Transfer*, **55**(6), 695–705.

Iati, M. A., R. Saija, A. Giusto, P. Denti, F. Borghese, C. Cecchi-Pestellini, 2004: Optical properties of interstellar gain aggregates, *J. Quant. Spectr. Rad. Transfer*, **89**, 43–45.

Ishimaru, A., 1978: *Wave Propagation and Scattering in Random Media* (2 vols), Academic Press, New-York.

Ishimaru, A., Y. Kuga, R. L. T. Cheung, K. Shimizu, 1983: Scattering and diffusion of a beam wave in randomly distributed scatterers, *J. Opt. Soc. Am.*, **73**(2), 131–136.

Jacquier, S., 2006: Approximate methods for the optical properties of spherical non-absorbent aggregated particles. PhD thesis, Ecole Nationale Supérieure des Mines de Saint-Etienne, Saint-Etienne.

Jacquier, S. and F. Gruy, 2007a: Approximation of the light scattering cross-section for aggregated spherical non-absorbent particles, *J. Quant. Spectr. Rad. Transfer*, **106**, 133–144.

Jacquier, S. and F. Gruy, 2007b: Approximation for scattering properties of aggregated spherical particles, PARTEC 2007, Nürnberg.

Jacquier, S. and F. Gruy, 2008a: Anomalous Diffraction Approximation for light scattering cross-section: case of ordered clusters of non-absorbent spheres, *J. Quant. Spectr. Rad. Transfer*, **109**, 789–810.

Jacquier, S. and F. Gruy, 2008b: Anomalous Diffraction Approximation for light scattering cross-section: case of random clusters of non-absorbent spheres, *J. Quant. Spectr. Rad. Transfer*, **109**, 2794–2803.

Kahnert, F. M., 2003: Numerical methods in electromagnetic scattering theory, *J. Quant. Spectr. Rad. Transfer*, **79**(80), 775–824.

Khlebtsov, N. G., 1996: Spectroturbidimetry of fractal clusters: test of density correlation function cut-off, *Appl. Optics*, **35**(21), 4261–4270.

Kimura, H. and I. Mann, 1998: Radiation pressure cross-section for fluffy aggregates, *J. Quant. Spectr. Rad. Transfer*, **60**(3), 425–438.

Kimura, H., H. Okamoto, T. Mukai, 2002: Radiation pressure and the Pointing-Robertson effect for fluffy dust particles, *Icarus*, **157**, 349–361.

Kokhanovsky, A. A., 2001: *Light Scattering Media Optics: Problems and Solutions* (2nd edn), Praxis Publishing, Chichester.

Kolokolova, L. and B. A. S. Gustafson, 2001: Scattering by inhomogeneous particles: microwave analog experiments and comparison to effective medium theories, *J. Quant. Spectr. Rad. Transfer*, **70**, 611–625.

Kostoglou M., A. G. Konstandopoulos, S. K. Friedlander, 2006: Bivariate population dynamics simulation of fractal aerosol aggregate coagulation and restructuring, *Aerosol Sci.*, **37**(9), 1102–1115.

Kruyt, H. R., 1952: *Colloid Science*, Elsevier, Amsterdam.

Kuga, Y. and A. Ishimaru, 1984: Retroreflectance from a dense distribution of spherical particles, *J. Opt. Soc. Am.*, **1**(8), 831–835.

Liu, C.-L., 1998: Validity of anomalous diffraction approximation in m-x domain, *Atmos. Res.*, **49**, 81–86.

Liu, Y., W. P. Arnott, J. Hallett, 1998: Anomalous diffraction theory for arbitrarily oriented finite circular cylinders and comparison with exact T-matrix results, *Appl. Optics*, **37**(21), 5019–5030.

Lopatin, V. N. and F. Ya. Sid'Ko, 1988: *Introduction to Optics of Cell Suspensions*, Moscow, Nauka.

Mekki-Berrada, K., F. Gruy, M. Cournil, 2005: *Synthèse d'agglomérats multi-échelles par précipitation homogène (Récents Progrès en Génie des Procédés)* Edition Lavoisier, Paris.

Mishchenko, M. I., D. W. Mackowski, L. D. Travis, 1995: Scattering of light by bispheres with touching and separated components, *Appl. Optics*, **34**(21), 4589–4599.

Mishchenko, M. I., L. D. Travis, and A. A. Lacis, 2002: *Scattering, Absorption, and Emission of Light by Small Particles*, Cambridge University Press, Cambridge.

Mishchenko, M. I., G. Videen, V. A. Babenko, N. G. Khlebtsov, T. Wriedt, 2004: T-matrix theory of electromagnetic scattering by particles and its applications: a comprehensive reference database, *J. Quant. Spectr. Rad. Transfer*, **88**, 357–406.

Mishchenko, M. I., G. Videen, V. A. Babenko, N. G. Khlebtsov, T. Wriedt, 2007: Comprehensive T-matrix reference database: A 2004–06 update, *J. Quant. Spectr. Rad. Transfer*, **106**, 304–324.

Mishchenko, M. I., G. Videen, N. G. Khlebtsov, T. Wriedt, N. T. Zakharova, 2008: Comprehensive T-matrix reference database: A 2006–07 update, *J. Quant. Spectr. Rad. Transfer*, **109**, 1447–1460.

Nichols, M. G., E. L. Hull, T. H. Forster, 1997: Design and testing of a white-light, steady-state diffuse reflectance spectrometer for determination of optical properties of highly scattering systems, *Appl. Optics*, **36**, 93–104.

Quirantes, A., F. Arroyo, et al., 2001: Multiple light scattering by spherical particle systems and its dependence on concentration; a T-matrix study, *J. Colloid Interf. Sci.*, **240**, 78–82.

Randolph, A. D., and M. A. Larson, 1988: *Theory of Particulate Processes*, Academic Press, New York.

Rouleau, F., 1996: Electromagnetic scattering by compact clusters of spheres, *Astron. Astrophys.*, **310**, 686–698.

Streekstra, G. J., A. G. Hoekstra, et al., 1994: Anomalous diffraction by arbitrarily oriented ellipsoids: applications in ektacytometry, *Appl. Optics*, **33**, 7288–7296.

Sugimoto, T., 2000: *Fine Particles: Synthesis, Characterization, and Mechanisms of Growth*, (Surfactant Science Series, Vol. 92), Marcel Dekker, New York.

Sun, W. and Q. Fu, 1999: Anomalous diffraction theory for arbitrarily oriented hexagonal crystals, *J. Quant. Spectr. Rad. Transfer*, **63**, 727–737.

Sun, W. and Q. Fu, 2001: Anomalous diffraction theory for randomly oriented nonspherical particles: a comparison between original and simplified solutions, *J. Quant. Spectr. Rad. Transfer*, **70**, 737–747.

Tandon, P., D. E. Rosner, 1999: Monte Carlo Simulation of Particle Aggregation and Simultaneous Restructuring, *J. Colloid and Interf. Sci.*, **213**, 273–286.

Tontrup, C., F. Gruy and M. Cournil, 2000: Turbulent aggregation of titania in water, *J. Colloid and Interf. Sci.*, **229**, 511–525.

Tsang, L., and A. Ishimaru, 1984: Backscattering enhancement of random discrete scatterers, *J. Opt. Soc. Am.*, **1**, 836–839.

Tsang L., and A. Ishimaru, 1985: Theory of backscattering enhancement of random discrete isotropic scatterers based on the summation of all ladder and cyclical terms, *J. Opt. Soc. Am. A*, **2**, 1331–1338.

Van de Hulst, H. C., 1981: *Light Scattering by Small Particles*, Dover publications Inc., New York.

Videen, G. and P. Chylek, 1998: Anomalous diffraction approximation limits, *Atmos. Res.*, **49**, 77–80.

Voshchinnikov, N. V. and V. G. Farafonov, 1992: Optical properties of spheroidal particles, *Astrophys. Space Sci.*, **204**, 19–86.

Wolf, P. E., G. Maret, E. Akkermans, R. Maynard, 1988: Optical coherent backscattering by random media: an experimental study, *J. Phys. Fr.*, **49**, 63–75.

Wriedt, T., 1998: A review of elastic light scattering theories, *Part. Part. Syst. Charact.*, **15**, 67–74.

Xu, Y.-L., 1995: Electromagnetic scattering by an aggregate of spheres, *Appl. Optics*, **34**(21), 4573–4588.

Xu, Y.-L., 1996: Calculation of the addition coefficients in electromagnetic multisphere-scattering theory, *J. Comput. Phys.*, **127**, 285–298.

Xu, Y.-L., 1997a: Electromagnetic scattering by an aggregate of spheres: far field, *Appl. Optics*, **36**(36), 9496–9508.

Xu, Y.-L., 1997b: Fast evaluation of gaunt coefficients: recursive approach, *J. Comput. Appl. Math.*, **85**, 53–65.

Xu, Y.-L., 1998a: Efficient evaluation of vector translation coefficients in multiparticle light-scattering theories, *J. Comput. Phys.*, **139**, 137–165.

Xu, Y.-L., 1998b: Electromagnetic scattering by an aggregate of spheres: asymmetry parameter, *Phys. Lett. A*, **249**, 30–36

Xu, Y.-L. and B. A. S. Gustafson, 2001: A generalized multiparticle Mie-solution: further experimental verification, *J. Quant. Spectr. Rad. Transfer*, **70**, 395–419.

Yang, P. and K. Liou, 2000: Finite difference time domain method for light scattering by nonspherical and inhomogeneous particles. In *Light Scattering by Nonspherical Particles: Theory, Measurements, and Applications*, eds. M. I. Mishchenko, J. W. Hovenier, and L. D. Travis, Academic Press, San Diego.

Yang, P., Z. Zhang, B. A. Baum, H. L. Huang, Y. X. Hu, 2004: A new look at anomalous diffraction theory (ADT): Algorithm in cumulative projected-area distribution domain and modified ADT, *J. Quant. Spectr. Rad. Transfer*, **89**, 421–442.

Part II

Modern Methods in Radiative Transfer

3 Using a 3-D radiative transfer Monte–Carlo model to assess radiative effects on polarized reflectances above cloud scenes

C. Cornet, L. C-Labonnote, and F. Szczap

3.1 Introduction

In the near future, more and more spaceborne or airborne instruments will be able to measure polarized reflectance issued from the atmosphere. To give some examples, currently, the POLarization and Directionality of the Earth's Reflectance instrument POLDER3/ PARASOL, which is the successor of POLDER2/ADEOS2 and POLDER/ADEOS (Deschamps et al., 1994) measures, since 2005, the polarized signal in the visible spectral range with up to 14 viewing directions. The airborne version of this instrument, called OSIRIS (observing system including polarization in the solar infrared spectrum (Auriol et al., 2008)), is nowadays extended to the near-infrared range and will maybe, in the future, generate a spaceborne version. The Aerosol Polarimetry Sensor (APS), the spaceborne version of the Research Scanning Radiometer (RSP) will be able to measure reflected total and polarized light in visible, near infrared, and short-wave infrared and should be launched in the framework of the Glory mission in 2010 (Mishchenko et al., 2007).

Measurements of polarized reflectances allow better retrieval of cloud and aerosol properties. Indeed, multi-angular polarimetric data can, for instance, be used to retrieve cloud phase (Gouloub et al., 2000; Riedi et al., 2001), cloud particle size (Bréon and Goloub, 1998; Bréon and Doutriaux-Boucher, 2005) or information on cloud particle shape (Chepfer et al., 2001; Herman et al., 2005; Sun et al., 2006). To retrieve these atmospheric properties accurately, or at least to assess the error which can be made with the classical retrieval algorithm, it is important to be able to model radiative transfer in a realistic way. Today, most of the code handling the polarization of light is based on the assumption of a plane-parallel atmosphere (de Haan et al., 1987; Stamnes et al., 1988; Rozanov and Kokhanovsky, 2006; Lenoble et al., 2007). However, it had been shown that this assumption can lead to non-negligible error in radiative quantities, especially in cloudy atmosphere (Marshak and Davis, 2005) and thus in the retrieved cloud properties (Loeb and Coakley, 1998; Varnai and Marshak, 2001; Iwabuchi and Hayasaka, 2002; Zinner and Mayer, 2006; Marshak et al., 2006; Cornet and Davies, 2008). These different studies deal with total reflectances and, so far, nobody has looked at the effects of cloud variabilities on polarized reflectances. The first step for doing these studies is to develop a model able to simulate polarized reflectances issued from a 3-D

atmosphere. Although polarization has been included recently in the Monte–Carlo code MYSTIC (Emde et al., 2009), most of the 3-D radiative transfer models existing now, either based on an explicit representation of the computations (Evans, 1998; Ferlay and Isaka, 2006) or on a statistical representation as are the models using the Monte-Carlo methods (Iwabuchi, 2006; Mayer, 2009), compute only total reflectances. Note that concerning the differences between the two methods, a comparative study between SHDOM (Evans, 1998) and one Monte–Carlo model can be found in Pincus and Evans (2009).

In order to study 3-D cloud radiative effects, we have also recently developed a three-dimensional and polarized radiative transfer Monte Carlo model called 3DMCpol (Cornet et al., 2009). This code is based on forward Monte–Carlo techniques and performs monochromatic calculations limited to the solar wavelength region: scattering and absorption processes are included, but not emission. It computes the complete Stokes vector (I,Q,U,V) in a 3-D atmosphere composed of cloud, aerosols and molecules for different view zenith and azimuth angles and hence allows the computation of the total and polarized reflectances. In this chapter, we present the model with additional tests concerning its validity and show two examples concerning the effects of cloud variabilities on polarized radiances. The chapter is organized as follows. After recalling the definitions concerning polarized light, the second section presents succinctly the polarized radiative transfer model, 3DMCpol (Cornet et al., 2009). Comparisons with a 1-D radiative transfer model are presented in section 3.3. In section 3.4, we show results of 3-D simulations obtained from 3-D synthetic clouds. Conclusions are given in section 3.5.

3.2 Including the polarization in a 3-D Monte–Carlo atmospheric radiative transfer model

3.2.1 Description of radiation and single scattering: Stokes vector and phase matrix

This section provides the description and the definition required to understand polarization of an electromagnetic radiation and how it can be modified by a scattering event. More details can be found in Hansen and Travis (1974) and Van de Hulst (1980).

3.2.1.1 Description of a beam of light

Light cannot be described by its intensity alone. Due to its wave nature, a complete description of light is obtained by taking into account its polarization state. Indeed, the radiation usually measured is often partially polarized and a beam of light consists of a polarized and an unpolarized part:

$$I = I_{pol} + I_{unpol} \,. \tag{3.1}$$

As the electric field describing the polarization of light draw an ellipse, the polarization part of the light can moreover be separated into a linearly polarized

part I_{lp} and a circularly polarized part I_{cp} as:

$$I_{pol} = \sqrt{I_{lp}^2 + I_{cp}^2}\,. \tag{3.2}$$

A beam of light is totally polarized only if the orientation and the ellipticity of the polarization ellipse is the same for each simple wave. On the other hand, light composed of uncorrelated waves and having no preferred direction for its electric field such as solar radiation is unpolarized.

We consider a parallel beam of light of wavelength λ traveling in the positive direction $\boldsymbol{z}$ and being scattered by particles of any shape. The scatterer is in the cartesian coordinate system ($\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$) attached to the photon with origin in the center of the scatterer. Note that the light is scattered in all directions of the space but for simplicity we will focus and describe scattering process in the direction defined by the vector $\boldsymbol{e_r}$. The plane defined by vectors $\boldsymbol{e_r}$ and $\boldsymbol{z}$ is called the scattering plane and is used, in the following, as the reference plane. The angle Θ between directions $\boldsymbol{e_r}$ and $\boldsymbol{z}$ is the zenith scattering angle. The scattering azimuth angle ϕ is the angle between the scattering plane and the $\boldsymbol{x}$ axis. The complex components of the electric field, parallel and perpendicular to the scattering plane are noted respectively $\boldsymbol{E}_{\parallel}$ and $\boldsymbol{E}_{\perp}$. The incident electric field can then be written as:

$$\boldsymbol{E_i} = \boldsymbol{E_{i\parallel}} + \boldsymbol{E_{i\perp}} = E_{i\parallel}\boldsymbol{e}_{\parallel} + E_{i\perp}\boldsymbol{e}_{\perp}\,, \tag{3.3}$$

where $\boldsymbol{e}_{\parallel}$ and $\boldsymbol{e}_{\perp}$ are the unit vector parallel and perpendicular to the scattering plane and such as $\boldsymbol{e}_{\parallel} \times \boldsymbol{e}_{\perp} = \boldsymbol{z}$.

The scattering wave is a spherical wave with an amplitude inversely proportional to the distance r. Moreover, as only linear processes happen when the incident field is modified in the scattered field, we can write (Van de Hulst, 1980):

$$\begin{pmatrix} E_{s\parallel} \\ E_{s\perp} \end{pmatrix} = \frac{e^{-ikr+ikz}}{ikr} \begin{pmatrix} A_2 & A_3 \\ A_4 & A_1 \end{pmatrix} \begin{pmatrix} E_{i\parallel} \\ E_{i\perp} \end{pmatrix}, \tag{3.4}$$

where k is the wavenumber and z the location.

This equation shows that scattering by particles, in every directions, is completely described by its amplitude matrix A. The elements of this matrix are complex and are function of the optical properties of the scatterer and of the scattering angles Θ, ϕ.

3.2.1.2 The Stokes parameters

The complete description of the polarization state of electromagnetic radiation can be obtained by the Stokes parameters. These parameters are the quantities computed in the radiative transfer model described in this chapter. If we consider a plane wave with parallel and perpendicular components (Chandrasekar, 1960), the Stokes parameters are (Van de Hulst, 1980):

$$\begin{cases} I = E_{\parallel}E_{\parallel}^* + E_{\perp}E_{\perp}^* \\ Q = E_{\parallel}E_{\parallel}^* - E_{\perp}E_{\perp}^* \\ U = E_{\parallel}E_{\perp}^* + E_{\perp}E_{\parallel}^* \\ V = i(E_{\parallel}E_{\perp}^* - E_{\perp}E_{\parallel}^*) \end{cases} \tag{3.5}$$

The asterisk stands for the complex conjugate. The component parallel and perpendicular of the electric field are described by their amplitudes $a_{\parallel}$ and $a_{\perp}$ and their phases $\epsilon_{\parallel}$ and $\epsilon_{\perp}$ chosen at ($z = 0$, $t = 0$) by:

$$\begin{array}{l} E_{\parallel} = a_{\parallel} e^{-i\epsilon_{\parallel}} e^{-ikz+i\omega t}\,, \\ E_{\perp} = a_{\perp} e^{-i\epsilon_{\perp}} e^{-ikz+i\omega t}\,, \end{array} \tag{3.6}$$

where ω is the frequency of the wave, k the wavenumber, z the location and t the time. From this expression, equation (3.5) becomes:

$$\begin{cases} I = a_{\parallel}^2 + a_{\perp}^2 \\ Q = a_{\parallel}^2 - a_{\perp}^2 \\ U = 2a_{\parallel} a_{\perp} \cos\delta \\ V = 2a_{\parallel} a_{\perp} \sin\delta \end{cases} \tag{3.7}$$

where $\delta = \epsilon_{\parallel} - \epsilon_{\perp}$.

I represents the intensity or normalized total reflectance. The other parameters having the same dimension, it allows them to be summed for a mixture of separate independent waves. Note that for clarity reasons, a constant factor common to all four parameters is omitted in these equations.

The Stokes parameters can also be represented by a geometric description. Indeed, the endpoints of the electric field vector draw an ellipse with semi-major axis and semi-minor axis called a and b respectively. The ellipticity is defined by an angle β such that:

$$\tan\beta = \pm\frac{b}{a}\,, \tag{3.8}$$

with $-\pi/4 \leq \beta \leq \pi/4$.

The sign of β gives the direction in which the ellipse is described, with positive sign corresponding to right-handed polarization. The angle between the major axis and the parallel direction ($\boldsymbol{e}_{\parallel}$) of the electric field is denoted by χ ($0 \leq \chi \leq \pi$).

The Stokes parameters can then be expressed by the parameters β and χ which define the shape and the orientation of the ellipse (Fig. 3.1; Van de Hulst, 1980) by:

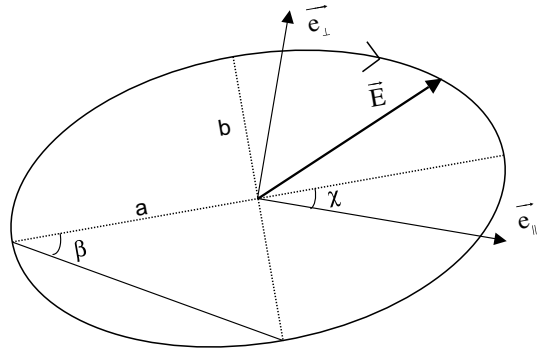


Fig. 3.1. Polarization ellipse. a and b are respectively the semi-major and the semi-minor axis. The direction of propagation is into the page and the sense of polarization indicated corresponds to right-handed polarization.

$$\begin{aligned} I &= S_0 \\ Q &= S_0 p \cos 2\beta \cos 2\chi \\ U &= S_0 p \cos 2\beta \sin 2\chi \\ V &= S_0 p \sin 2\beta \end{aligned} \tag{3.9}$$

where S_0 is the total intensity of the beam of light and p the degree of polarization.

From this equation, it is possible to find characteristics of the polarization ellipse for the Stokes parameters by:

$$\begin{cases} \beta = \dfrac{1}{2} \arctan\left(\dfrac{V}{\sqrt{Q^2+U^2}}\right) \\ \chi = \dfrac{1}{2} \arctan\left(\dfrac{U}{Q}\right) \end{cases} \tag{3.10}$$

The exact relations between $a_\parallel$, $a_\perp$, δ in equation (3.7) and S_0, β, and χ in equation (3.9) can be found in Van de Hulst (1980).

Another way to understand the meaning of the Stokes parameters consists of looking at their measurements. If we suppose that $\boldsymbol{E}_\parallel$ has a constant delay ε with respect to $\boldsymbol{E}_\perp$ and that $I(\psi, \varepsilon)$ is the intensity of light due to vibrations in the direction making an angle ψ with $\boldsymbol{e}_\parallel$. The Stokes parameters can be obtained from (Hansen and Travis, 1974):

$$\begin{aligned} I &= I(0°, 0) + I(90°, 0) = I_\parallel + I_\perp \\ Q &= I(0°, 0) - I(90°, 0) = I_\parallel - I_\perp \\ U &= I(45°, 0) - I(135°, 0) \\ V &= I(45°, \tfrac{\pi}{2}) - I(135°, \tfrac{\pi}{2}) \end{aligned} \tag{3.11}$$

I represents the total intensity. Q is the difference between the intensity of light measured with a polarizer in the parallel direction with respect to the reference plane ($\psi = 0°$) and the intensity measured in the perpendicular direction ($\psi = 90°$). U has a similar definition to Q but the difference is between intensity for $\psi = 45°$ and for $\psi = 135°$. Finally, V is the difference between a right-handed circular polarization and a left-handed circular polarization.

For monochromatic light, which is fully polarized, we obtain from equation (3.7):

$$I^2 = Q^2 + U^2 + V^2 \tag{3.12}$$

For natural light, because this is a mixture of uncorrelated simple waves with no preferred direction of vibration, there exists no phase relation between the parallel and perpendicular fields. Thus, we have

$$Q = U = V = 0 \tag{3.13}$$

Light is thus unpolarized and I is sufficient to describe it. But usually, the light is partially polarized and :

$$I^2 \geq Q^2 + U^2 + V^2 \tag{3.14}$$

As already mentioned, a beam of light can thus be considered as the sum of polarized and unpolarized light (equation 3.1) with the intensity of polarized light

equal to:

$$I_{pol} = (Q^2 + U^2 + V^2)^{\frac{1}{2}} \tag{3.15}$$

From the Stokes parameters, the following quantities are also defined:

- the degree of polarization (equation (3.9)): $p = \frac{\sqrt{Q^2+U^2+V^2}}{I}$
- the intensity of linearly polarized light $\sqrt{Q^2 + U^2}$
- the intensity of circularly polarized light: V

3.2.1.3 The phase matrix

We can see in equation (3.4) that the scattered electric field is obtained from the incident electric field by a simple multiplication with the amplitude matrix of the scatterer. When working with the Stokes parameters, we use in the same way the phase matrix $P(\Theta)$ which allows us to transform the incident Stokes vector in the scattered Stokes vector:

$$\begin{pmatrix} I_s \\ Q_s \\ U_s \\ V_s \end{pmatrix} = \frac{1}{k^2 R^2} \begin{pmatrix} P_{11} & P_{12} & P_{13} & P_{14} \\ P_{21} & P_{22} & P_{23} & P_{24} \\ P_{31} & P_{32} & P_{33} & P_{34} \\ P_{41} & P_{42} & P_{43} & P_{44} \end{pmatrix} \begin{pmatrix} I_i \\ Q_i \\ U_i \\ V_i \end{pmatrix}, \tag{3.16}$$

where $k = 2\pi/\lambda$ and R is the distance between the observer and the particles. As only linear processes happen, it is possible to express the elements of the phase matrix as functions of the elements of the amplitude matrix (Van de Hulst, 1980). In the phase matrix, note that P_{11} corresponds to the phase function used in the scalar approach and describes the probability of a photon being scattered along the direction defined by Θ with respect to the incident direction. It is normalized as:

$$\int_0^{2\pi} \int_0^1 P_{11}(\Theta)\, d\cos(\Theta)\, d\varphi = 4\pi \tag{3.17}$$

Under some assumptions, the phase matrix of a group of particles can be expressed as the sum of each individual phase matrix. Consequently, Van de Hulst (1980) showed thus that a group of particles leads to symmetry relationships which allow the simplification of the phase matrix of the ensemble.

For example, a cloud composed of randomly oriented particles with no optically active sphere leads to a phase matrix with a maximum of six independent parameters function of Θ, and can be written as:

$$\mathbf{P} = \begin{pmatrix} P_{11} & P_{12} & 0 & 0 \\ P_{12} & P_{22} & 0 & 0 \\ 0 & 0 & P_{33} & P_{34} \\ 0 & 0 & -P_{34} & P_{44} \end{pmatrix}. \tag{3.18}$$

In addition, for spherical particles, we have:

$$\begin{cases} P_{11}(\theta) = P_{22}(\theta) \\ P_{33}(\theta) = P_{44}(\theta) \\ P_{11}^2(\theta) - P_{33}^2(\theta) - P_{12}^2(\theta) - P_{34}^2(\theta) = 0 \end{cases} \tag{3.19}$$

In Fig. 3.2, we have plotted the elements of the phase matrix at a wavelength of 865 nm for three types of scattering:

- Rayleigh scattering for small particles, which is principally used to model gaseous molecular scattering.
- Mie scattering used for spherical particles such as water cloud droplets. The phase matrix is computed from a log-normal distribution with an effective radius of 10 μm and an effective variance of 0.02. The real part of the refractive index is 1.33 and the imaginary part is null.
- Scattering by non-spherical particles such as ice particles obtained with the IHM particles (Inhomogeneous Hexagonal Monocrystal model (C-Labonnote et al., 2001)). This model consists in computing the scattering of light by an ensemble of randomly oriented hexagonal ice crystals containing spherical impurities of air bubbles. The real part of the refractive index is 1.304 and the imaginary part is null.

Concerning Rayleigh scattering, the evolution of the phase matrix components as functions of the scattering angle is very smooth with important side scattering. They are simply expressed by:

$$\mathbf{P}_{ray}(\Theta) = \frac{3}{4}\begin{pmatrix} 1+\cos^2\Theta & -\sin^2\Theta & 0 & 0 \\ -\sin^2\Theta & 1+\cos^2\Theta & 0 & 0 \\ 0 & 0 & 2\cos\Theta & 0 \\ 0 & 0 & 0 & 2\cos\Theta \end{pmatrix} \tag{3.20}$$

Concerning spherical particles, we note that the relations expressed in equation (3.19) are respected. Another important characteristic is the variation of the parameter P_{12} for Θ larger than 130°. The large maxima near 140° corresponds to the so-called primary cloudbow and is characteristic of water cloud. It is used among others tests for cloud phase detection (Riedi et al., 2001). In addition, the positions of the secondary maximas and minimas between 145°and 165° are used to retrieve size parameters of cloud particles (Bréon and Doutriaux-Boucher, 2005). Note that the presence of these secondary cloudbows are only visible for relatively narrow size distribution. Here, the effective variance of the distribution is 0.02 for an effective radius of 10 μm.

For heterogeneous non-spherical particles, because of irregularities (inclusion or surface roughness), side scattering is larger comparing to liquid spherical particles and thus the assymetry parameter is smaller. Moreover, note that $P_{22} \neq P_{11}$ and $P_{44} \neq P_{33}$. The ice phase function (IHM) chosen here is just an example and does not reflect the diversity of ice crystal found in cirrus cloud. Indeed, depending of the ice crystal shape, the elements of the phase matrix can have very different behavior (Sun et al., 2006).

3.2.2 Description of the radiative transfer model, 3DMCpol

The radiative transfer model called 3DMCpol (Cornet et al., 2009) follows the standard approach of a forward Monte–Carlo model (Cashwell and Everett, 1959; Avery and House, 1969; Marchuk et al., 1980; Evans and Marshak, 2005; Iwabuchi,

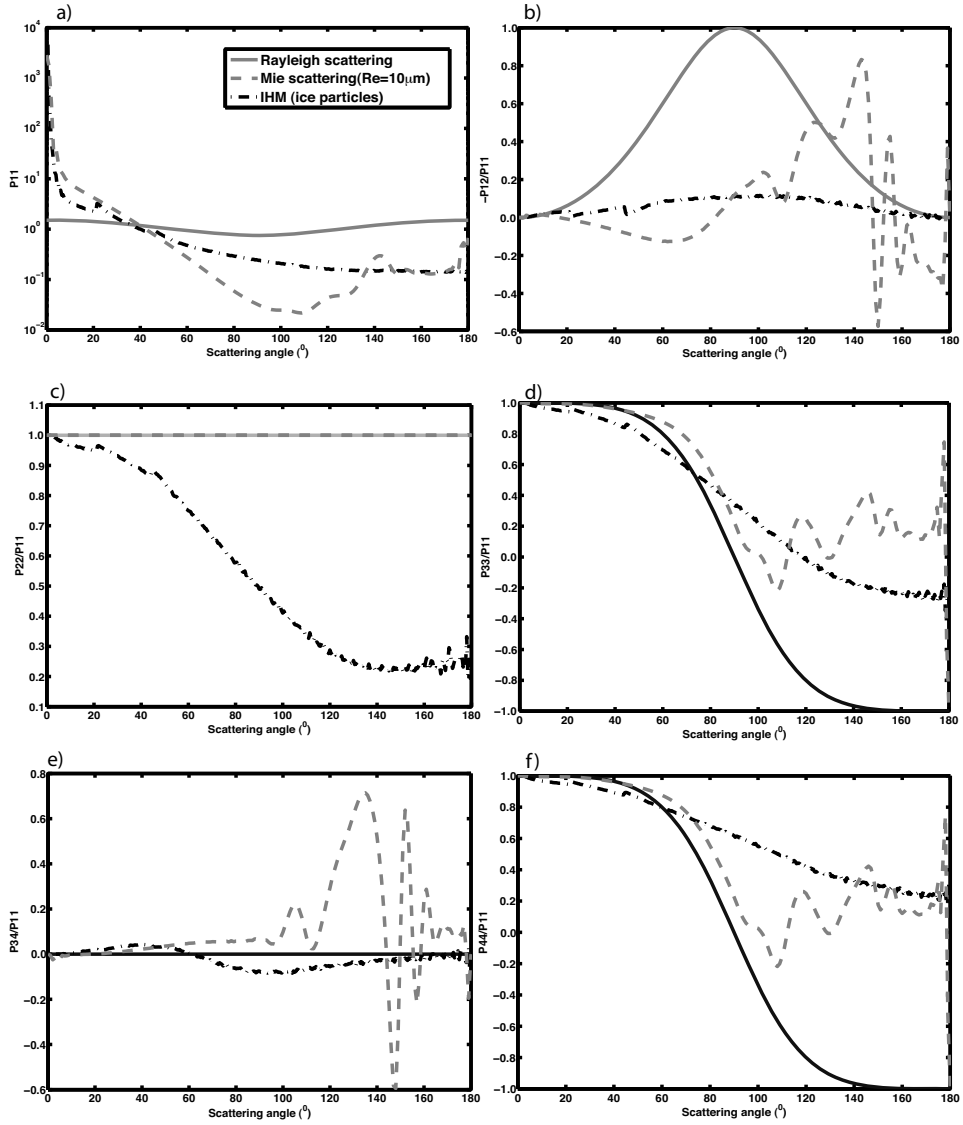


Fig. 3.2. Examples of scattering phase matrix components for a wavelength of 865 nm and for three types of scattering: Rayleigh scattering for small particles, Mie scattering for spherical particles ($R_{eff} = 10\,\mu$m and $V_{eff} = 0.02$) and randomly oriented non-spherical particles. The last is obtained with the IHM model (C-Labonnote et al., 2001).

2006; Mayer, 2009). A flowchart describing the progression for each photon is presented in Fig. 3.3. The output of the model is the normalized reflectance in percent noted:

$$I = \frac{\pi.R}{\mu_0 F_0} \times 100\,, \tag{3.21}$$

where R is the reflected radiance in $\mathrm{W\,m^{-2}\,sr^{-1}}$ and $\mu_0 F_0$ the incoming solar flux.

Computation of polarized reflectances (vector approach) does not change the general scheme of the code. Instead of scalar quantities, e.g. radiances (I), the Stokes vector $S = (I, Q, U, V)$ that completely describe the polarized state of the light (Chandrasekar, 1960) is computed. The phase function is thus substituted by the 4×4 phase matrix $P(\Theta)$. Otherwise, two important differences exist: one concerns the computation of the scattering azimuth angle and the other concerns the rotation of the electric field characterizing the polarization.

Indeed, first, the scattering azimuth angle for each scattering event, depends on the scattering zenith angle Θ and on the state of polarization of the incident photon. The uniform PDF of ϕ is replaced by a conditional PDF $f(\phi|\Theta)$, which is the probability of ϕ given Θ, whereas in the scalar approach, this angle is chosen randomly between $-\pi$ and π. The azimuth angle ϕ is found by resolving the following expression (more details can be found in Cornet et al. (2009)):

$$2\pi\zeta_\phi = \phi - \left(\frac{\sin 2\phi}{2} Q_i + U_i \sin^2\phi\right) D_p(\Theta)\,, \tag{3.22}$$

where ζ_ϕ is a uniform random number between 0 and 1.

Concerning the rotation of the Stokes vector S, because it is defined with respect to the meridian plane whereas the phase matrix is defined in the scattering plane, two rotations are needed. Therefore, in order to deal with a scattering event from the incident direction defined by (θ, φ) to a new direction defined by (θ', φ'), the incident Stokes vector S_i must be multiplied by the matrix Z (Chandrasekar, 1960; Hansen and Travis, 1974; Ramella-Roman et al., 2005):

$$Z(\theta, \theta', \varphi - \varphi') = R(\pi - i_2) P(\Theta) R(-i_1)\,, \tag{3.23}$$

where Θ is the zenith scattering angle and i_1 and i_2 are two rotation angles defined as follows (Hovenier, 1969; Hansen and Travis, 1974):

$$\begin{aligned} \cos i_1 &= \cos\phi \\ \cos i_2 &= \frac{-\cos\theta + \cos\theta' \cos\Theta}{\pm \sin\Theta \sin\theta'}\,, \end{aligned} \tag{3.24}$$

where ϕ is the scattering azimuth angle. The sign $\pm$ depends of the sign of the relative azimuth angle $(\varphi - \varphi')$.

The rotation matrix $\mathbf{R}$ is given by:

$$\mathbf{R}(i) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2i & \sin 2i & 0 \\ 0 & -\sin 2i & \cos 2i & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.25}$$

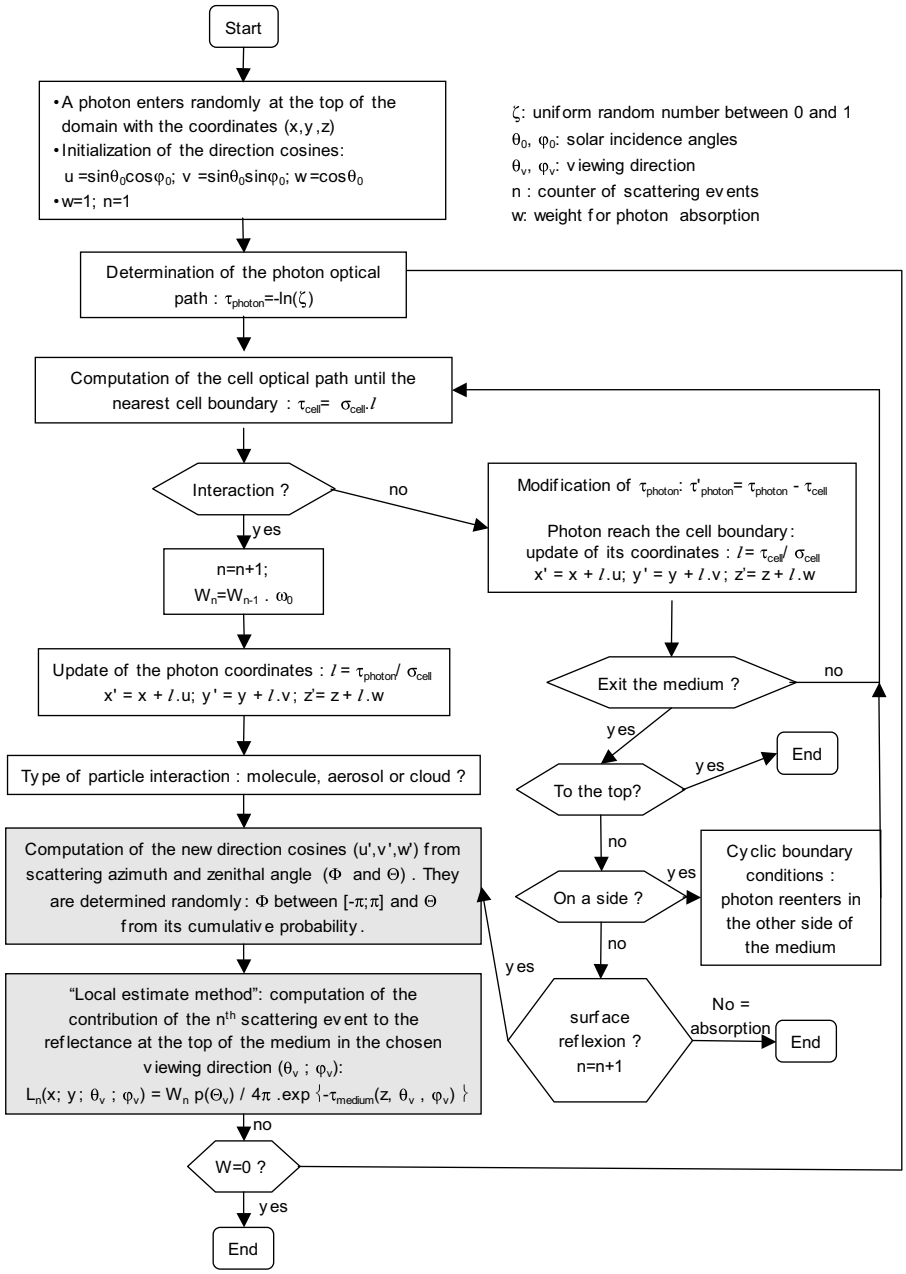


Fig. 3.3. Flowchart of the standard Monte–Carlo radiative transfer model for the progression of one photon. The gray boxes correspond to the steps where the principal modifications of the code take place in order to account for the state of polarization of the light.

Note that, so far, the polarized 3-D model used here allows only the computation of reflectances from a medium composed of randomly oriented particles and it does not handle the polarization in the case of surface reflection. Some modifications will be included later in order to allow computation of oriented particles and polarized surface reflection.

3.3 Total and polarized reflectances in the case of homogeneous clouds (1-D)

When the work described in this chapter was done, no 3-D atmospheric models accounting for the state of polarization of the light in the visible part of the spectrum were available. The validation of 3DMCPOL directly with 3-D polarized cases was therefore not possible. Consequently, the model was validated, on the one hand, for total reflectances in 3-D cases (see one example in subsection 3.4.2) and on the other hand, for polarized reflectances from 1-D homogeneous cloud (next subsection).

3.3.1 Validation of the MC polarized model

In this subsection, several comparisons made with the well-known Adding-Doubling code (de Haan et al., 1987) are shown for 1-D homogeneous cases for total and polarized reflectances. Concerning the Monte–Carlo simulations, 5.10^6 photons were launched.

The first comparison concerns Rayleigh scattering which is one of the simplest kinds of scattering. This case is a good validation for our code, first, because the phase matrix and its exact decomposition in Legendre polynomials used in the Adding-Doubling code is numerically known and, secondly, because Rayleigh scattering is known to highly polarize the light. Results are shown in Fig. 3.4. Reflectances are computed for a cloud with optical thickness 5 using the Rayleigh phase matrix, for different view zenith angles (0°, 30° and 60°) and for different view azimuth angles between 0° (backward direction) and 180° (forward direction). The sun incidence is 60°. Note that we only consider the cloud layer which means that, in this example, atmospheric molecular scattering was not accounted for. As the last component V is zero for Rayleigh scattering, we present results for the first three components of the Stokes vector (I,Q,U) respectively in Fig. 3.4(a), (c) and (e). Comparisons obtained for the polarized reflectance I_p that is commonly used in retrieval algorithms are also plotted in Fig. 3.4(g). This last quantity describes the amount of polarized light:

$$I_p = \varepsilon\sqrt{Q^2 + U^2 + V^2} \tag{3.26}$$

where $\varepsilon = \pm 1$ represents the sign of the polarized reflectances which is expressed in (C-Labonnote et al., 2001), but note that, in the solar plane, it corresponds simply to the sign of Q. In Fig. 3.4(b), (d), (f), (h), we report the relative difference in percent obtained between the two codes.

We see that, whatever the viewing angle, the agreement between the two models is very good for all the components of the Stokes vector as well as for the polarized

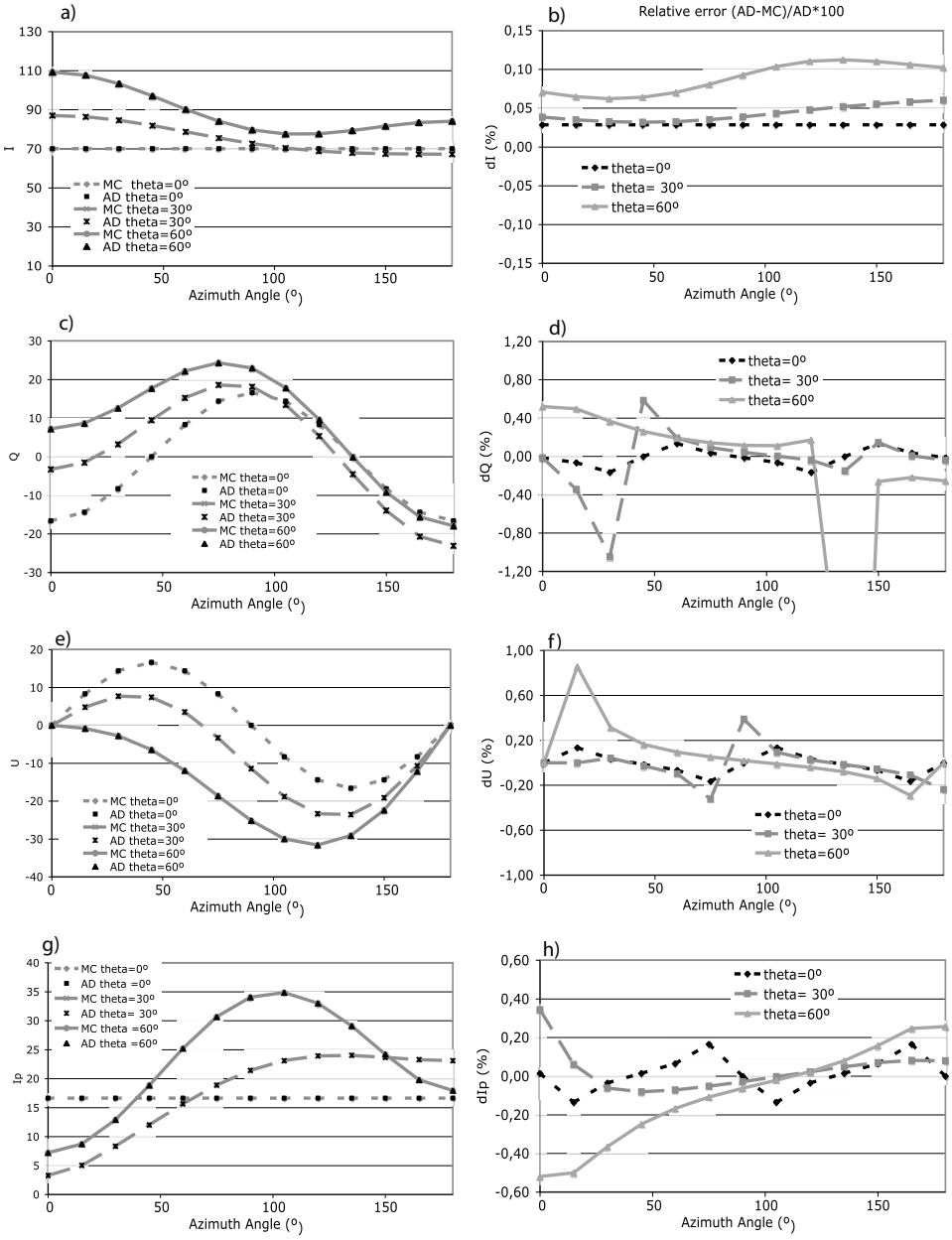


Fig. 3.4. Comparisons between Monte–Carlo and Adding-Doubling simulations for a homogeneous cloud with $\tau = 5$ using a Rayleigh phase matrix. $\theta_s = 60°$, $\theta_v = 0°$, 30° or 60° and φ_v between 0° (backward scattering) and 180° (forward). Figs (a), (c), (e), and (g) present results for the Stokes parameters respectively I, Q, U and for the polarized reflectances I_p. All these quantities are presented in percent. Figs (b), (d), (f), and (h) present the relative differences in percent for the same quantities.

reflectances. Except when the value is close to zero, the relative differences are very small: below 0.12% for total reflectances and below 0.6% for polarized reflectances.

Because it will be used further in the chapter, we present also results obtained for a homogeneous cirrus cloud in Fig. 3.5. The cirrus cloud has an optical thickness of 2 and is located between 7 and 9 km. In this example, we account for atmospheric molecular (Rayleigh) scattering computed for a wavelength of 865 nm between 0 and 20 km but not for atmospheric absorption. The phase matrix used in the simulation comes from the IHM (Inhomogeneous Hexagonal Monocrystal) model (C-Labonnote et al., 2001). The error is below ±1% for I and about ±4–5% for Q, U and I_p. The differences observed for Q, U and I_p can be explained, on the one hand, by the small values reached by the polarized reflectances. Note, for that matter, that the value of polarized reflectances is much lower in the case of ice scattering than in case of Rayleigh scattering. The second reason concerns the different treatments of the phase matrix used in the two models. In the Monte–Carlo model, the exact phase matrix is used whereas in the Adding-Doubling code, for computational time reasons concerning the zenithal integration, some approximation like delta-M scaling or truncation of the phase function are made that could introduce a-physical behavior in the computed Stokes vector.

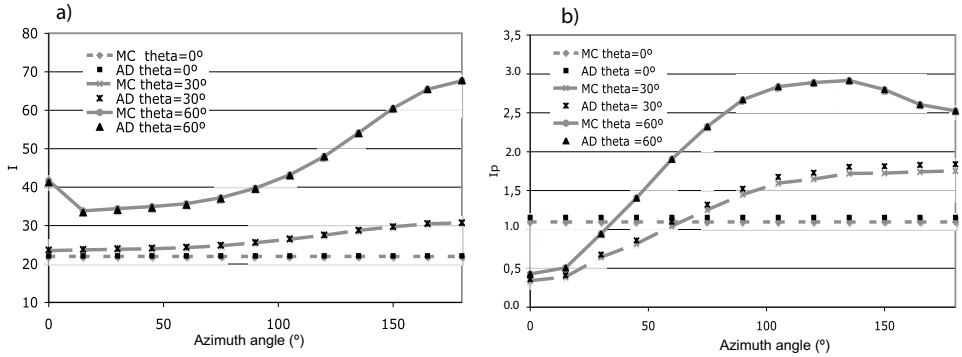


Fig. 3.5. Comparisons of the total reflectances (a) and the polarized reflectances (b) in percent of a homogeneous cirrus cloud with $\tau = 2$ computed with the Monte-Carlo code (MC) or with the Adding-Doubling code (AD). $\theta_s = 60°$, $\theta_v = 0°$, 30° or 60° and φ_v between 0° (backward) and 180° (forward). From Cornet et al. (2009).

The last comparison presented concerns a homogeneous water cloud with an optical thickness of 10 (Fig. 3.6). Again, we add molecular scattering for the atmosphere. The results are presented in the solar plane as a function of the scattering angle. In this case, the sign of the polarized reflectance corresponds to the sign of the second component of the Stokes vector Q as the third parameter U in the solar plane is null. The comparison shows that our model can well reproduce the primary cloudbow near 140° used for cloud phase detection and the secondary bows which are used to retrieve size particles of water cloud (Bréon and Goloub, 1998; Bréon and Doutriaux-Boucher, 2005). One can notice that the differences given by the two codes are very small whatever the direction. These differences are larger in the forward direction and are certainly due to the decomposition of the phase

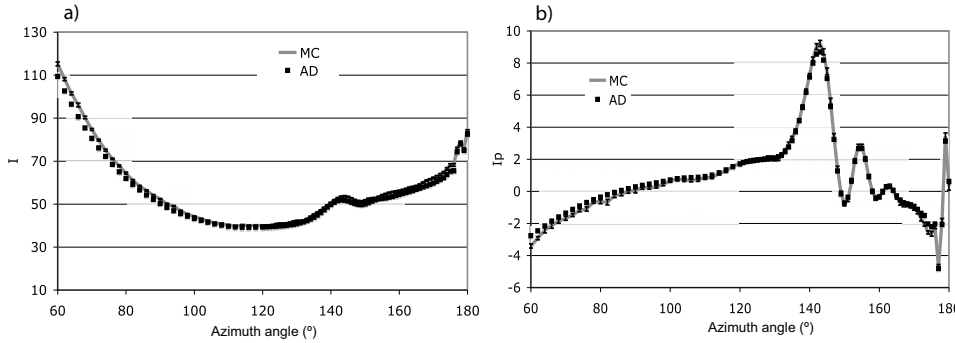


Fig. 3.6. Comparisons of the total reflectances (a) and the polarized reflectances (b) in percent of a homogeneous water cloud with $\tau = 10$ computed with the Monte–Carlo code (MC) or with the Adding-Doubling code (AD). $\theta_s = 60°$ and reflectances are plotted as functions of the scattering angle 0° (forward) and 180° (backward).

matrix in Legendre polynomials and to the number of integration points used in the Adding-Doubling code.

We recall that in this subsection, we have only shown comparisons on polarized reflectances obtained for 1-D homogeneous cloud. Examples of comparisons for total reflectances obtained for 3-D inhomogeneous cloud are presented in subsection 3.4.2.

3.3.2 Reflectances of homogeneous clouds as a function of the optical thickness

Fig. 3.7 presents, for homogeneous cirrus clouds, total and polarized reflectances as a function of optical thickness for a sun zenith angle of 60°, a zenith observation angle of 0° (a,b) and of 60° (c,d) and an azimuth angle of 0^0 (backward scattering direction). Under the homogeneous assumption (solid line), we recognize the well-known nonlinear relationship between reflectances and optical thickness. This nonlinear relationship exists for both total and polarized reflectances but, as already pointed out, we note that polarized reflectances saturate for relatively small optical thickness of about 3–4. This is because the final polarization state comes from the first few orders of scattering. This particularity is fundamental and makes the polarized measurements very useful in retrieving information on cloud particle shapes or size. Indeed, as polarized reflectances are less sensitive to multiple scattering, information contained in the measurement comes essentially from the particles' scattering properties via their phase matrix.

3.4 Total and polarized reflectances in the case of 3-D cloud fields

3.4.1 Description of the 3-D cloud fields used

In this subsection, we present computations of total and polarized reflectances obtained from 3-D clouds that are an inhomogeneous cirrus and a stratocumulus

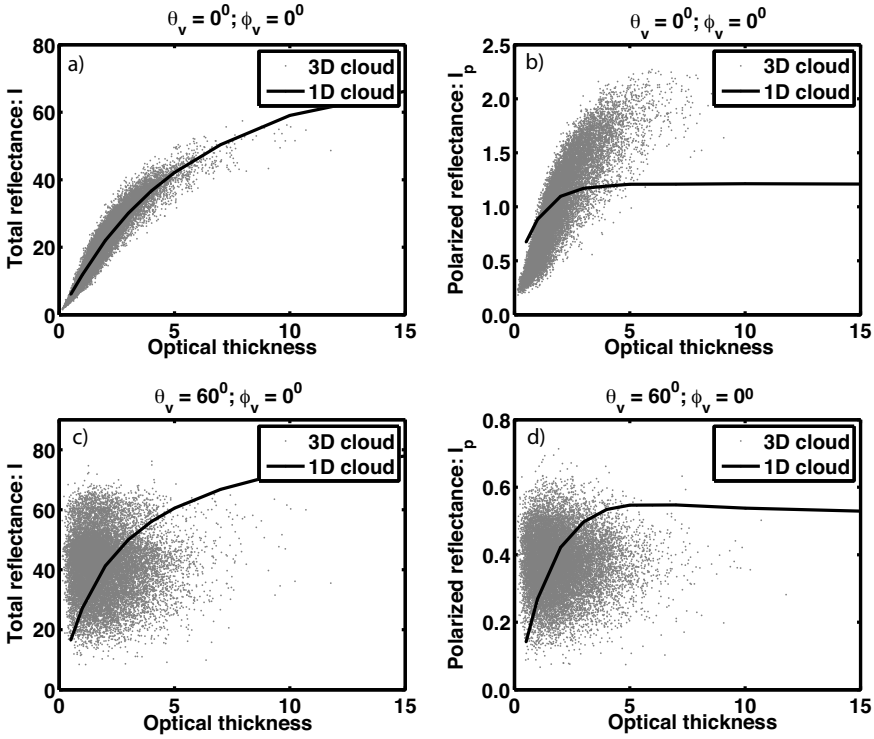


Fig. 3.7. Total (a,c) and polarized (b,d) reflectances in function of optical thickness computed for homogeneous cirrus clouds (solid line) and for the heterogeneous cirrus cloud presented in Fig. 3.8 (dot) with $\overline{\tau} = 2$. An IHM phase matrix is used. The sun zenith angle 60°, the azimuth observation angle is 0° and the zenith observation angle is 0° (a, b) or 60° (c,d).

cloud. The clouds were simulated with a stochastic cloud model called '3DCloud' and described in Szczap (2009) and Cornet et al. (2009). The simulations are done at a wavelength of 865 nm. The two cloud fields are composed of 128×128 pixels. The optical thickness of the cirrus cloud is a gamma PDF with a mean of 2, an inhomogeneity parameter ρ of 0.6, a fractional cloud cover of 1 and a wind shear of $2\,\mathrm{m\,s^{-1}}$. The inhomogeneity parameter is defined as the standard deviation of the 2-D horizontal optical depth fields divided by its mean (Szczap et al., 2000). The optical thickness of the cirrus cloud integrated along the z-axis is presented in Fig. 3.8(a) and the integration on the extinction coefficient along the y-axis in Fig. 3.8(b). The other cloud is a stratocumulus cloud with a mean optical thickness of 10, an inhomogeneity parameter of 0.7 and a fractional cloud cover of 1. It is presented in Fig. 3.8(c) and (d). As in subsection 3.3.1, the cirrus cloud microphysics is modelized by an IHM particles phase function computed with a real part of the refractive index set to 1.304. The stratocumulus cloud microphysics is modeled with spherical particles with a log-normal size distribution, an effective radius of $10\,\mu$m and an effective variance of 0.02. The real part of the refractive index is 1.33.

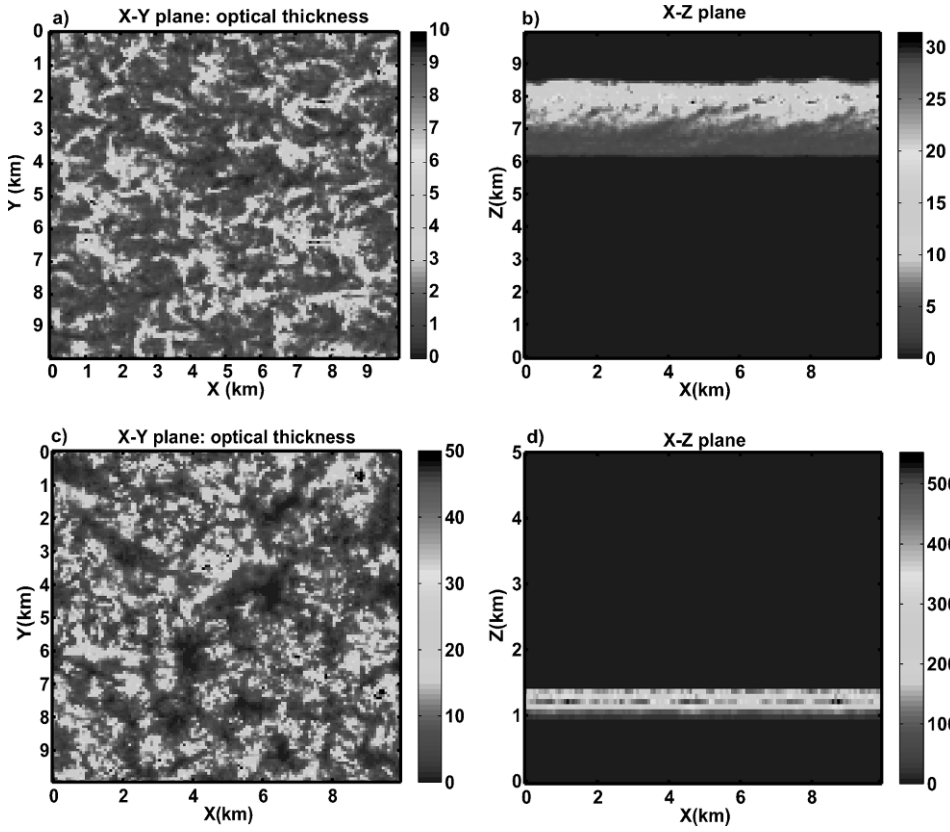


Fig. 3.8. Synthetic clouds used for the radiative transfer. (a) and (b) cirrus cloud with a mean optical thickness of 2. (c) and (d) stratocumulus cloud with a mean optical thickness of 10. Optical thickness integrated along the z-axis (a) and (c), and along the y-axis (b) and (d). (a) and (b) are from Cornet et al. (2009). [For the color version of this figure, see the color section.]

3.4.2 Comparisons with SHDOM and time considerations

In this subsection, we show comparisons obtained for the 3-D cirrus cloud presented in Fig. 3.8 with another widely used 3-D radiative transfer model, the Spherical Harmonics Discrete Ordinate Method (SHDOM, (Evans, 1998)). Radiative transfer calculations were performed for a solar zenith angle of 60. The medium is described by $128 \times 128 \times 44$ pixels with a size of 78 m along the x and y-axis. Along the z-axis, the pixel size is 78 m inside the cloud and 2 km outside the cloud to account for molecular scattering between 0 and 20 km. For all the calculations, the wavelength is 865 nm. We assume a black underlying surface. The results are presented in Fig. 3.9 for a zenith observation angle of 30° and a relative azimuth angle of 180^0 (forward direction). The simulations were performed with 10^8 photons divided in 20 batches which makes it possible to compute the statistical error of the Monte–Carlo calculations (Evans and Marshak, 2005). For this simulation, the absolute Monte–Carlo statistical errors for the total and polarized normalized reflectances

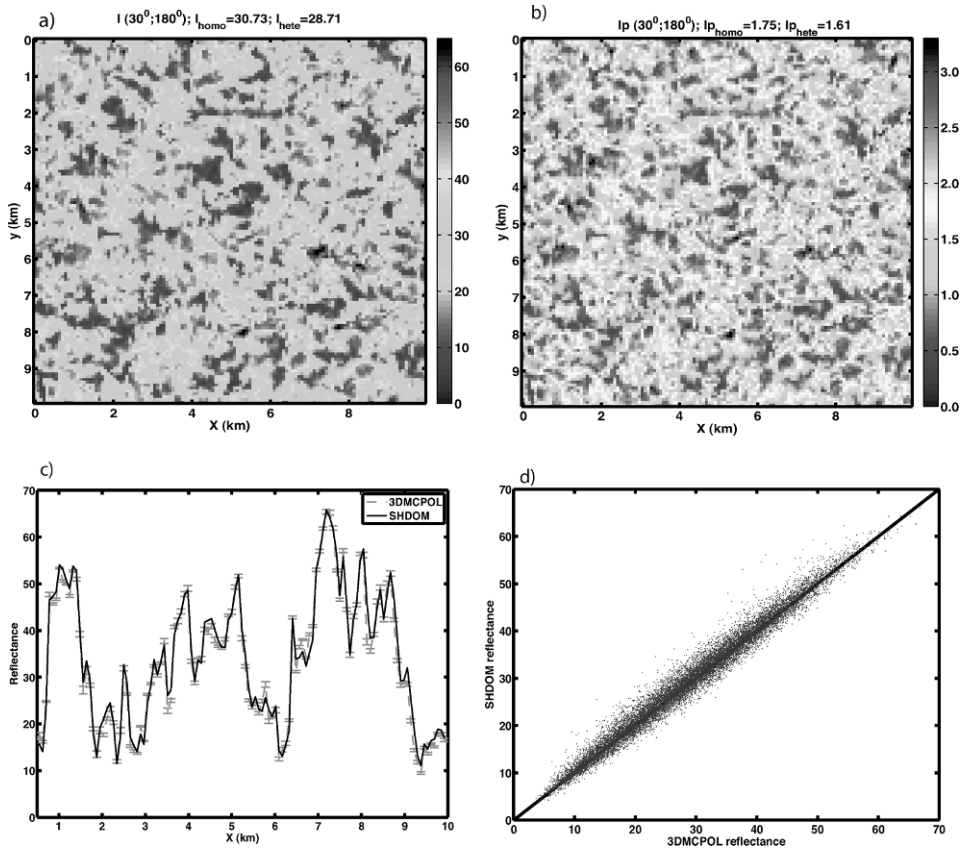


Fig. 3.9. Total (a) and polarized (b) reflectances at 78 m resolution for $\theta_s = 60°$; $\theta_v = 30$ and $\varphi_v = 180°$ (forward direction) for the cirrus cloud presented in Fig. (3.8). The wavelength is 865 nm. The mean values are indicated above the figures for the 3-D cirrus cloud (hete) and for the homogeneous cloud (homo). (c) and (d) comparisons with SHDOM for the total reflectance value: (c) horizontal cut along the line 75 corresponding to about 6 km and (d) scatterplot between SHDOM and 3DMCPOL reflectances for the entire field presented in (a). Error bars correspond to the statistical errors of Monte–Carlo simulations. (a) and (b) are from Cornet et al. (2009). [For the color version of this figure, see the color section.]

are respectively ± 1 and ± 0.12 which corresponds to relative statistical errors respectively below 5% and below 10% for the majority of the pixels. When averaging on the entire field, we obtain respectively statistical errors of 1.5% and 3%. Comparisons with SHDOM are presented in Fig. 3.9(c) and (d) where a horizontal cut along the line 75 and a scatterplot of the entire field are shown. We can see that Monte–Carlo and SHDOM results are in very good agreement with on average an absolute error of 0.5 corresponding to a relative error of 1.5%.

One of the main drawbacks of Monte–Carlo calculations concerns the large computer time which is required to reach a good accuracy. In our case, because we take into account the state of polarization of the light and all the computation

required for it, such as the determination of the scattering azimuth angle and the rotation of the Stokes vector, the simulations are slowed down by a factor of almost 2 compared to a model using the scalar approach. One way to reduce this computing time is to not compute the four elements of the Stokes vector when possible. Typically, atmospheric processes or interactions lead to linearly polarized light; therefore the fourth coefficient (V) of the Stokes vector is very small and can be neglected. Consequently we can compute just the first three elements of the Stokes vector which reduces the computation time by a factor of 1.2. This factor becomes 1.5 if we compute only the first two elements when calculations take place in the solar plane.

In subsection 3.3.1, we showed that the polarized Monte–Carlo model gives accurate results for polarized reflectances over different types of homogeneous cloud cases. In this section, we showed that the agreement is good for total reflectances over a 3-D cloud scene. Therefore, we can have confidence in results obtained for both total and polarized reflectances above 3-D scenes.

3.4.3 High spatial resolution (80 m): illumination and shadowing effects

Cloud heterogeneity effects can often be separated into two according to the scale being studied. In this subsection, we present results obtained for a high resolution of 80 m, which is close to what can be obtained by airborne measurements. Total and polarized reflectances are respectively presented in Fig. 3.9(a) and (b). We see that the two reflectances fields reproduce well the variability of the optical thickness fields. However, one can be surprised by the values reached by polarized reflectances. Indeed, as we saw in Fig. 3.7, in the case of homogeneous cloud, polarized signal saturates for optical thickness of about 3–4. This appears no more true in the case of variable 3-D cloud fields as it is confirmed in Fig. 3.7(b) and (d) where are plotted the reflectances obtained for this 3-D cloud as a function of the optical thickness. Both total and polarized 3-D reflectances are dispersed around the 1-D relationship and one can notice that the dispersion is more important for a zenith observation angle of 60° than for a zenith angle of 0°. Two effects can explained this dispersion. One is a well-known 3D effects already observed at high resolution on total reflectances (Varnai and Davies, 1999; Varnai and Marshak, 2002) and due to illumination and shadowing effects. This effect appears whatever the observation angle. It explained that, for a given optical thickness, reflectance values can be below or above the 1-D curve, which implies one part of the observed dispersion. For a zenith observation angle of 60°, another effect appears and comes from the fact that reflectances are the result of energy coming along the oblique path defined by the observation angle whereas the optical thickness is integrated along the vertical axis (Varnai and Davies, 1999). Regarding these results, we can conclude that at high resolution polarized reflectances are affected by 3-D effects in the same way as total reflectances. People need therefore to be cautious when using polarized measurements above variable cloud scenes because, in addition to erroneous retrieved values, illumination effects can lead to situations with no solution in the framework of homogeneous cloud layers. To better understand illumination and shadowing effects, one can refer to Cornet et al. (2009) where reflectances simulated from a step cloud are analyzed.

3.4.4 Medium spatial resolution (10 km): sub-pixel heterogeneity effects

To assess the sub-pixel heterogeneity effects at a larger resolution, we averaged spatially the total and polarized reflectance fields to obtain mean values over a 10×10 km pixel size, which is close to the POLDER pixel resolution. This spatial average of the total and polarized reflectances is plotted in Fig. 3.10 (dashed lines) for the 3-D cirrus cloud presented in Figs 3.8(a) and (b) and in Fig. 3.11 (dashed lines) for the 3-D stratocumulus cloud presented in Figs 3.8(c) and (d). In addition, we report also in these figures the reflectance obtained from a 1-D model, which is usually used in classical retrieval. We assume the same microphysical model and the averaged optical thickness over the 3-D domain. The cirrus cloud microphysics is thus represented by the IHM model with an optical thickness of 2 and the stratocumulus cloud by a log-normal distribution with $r_e = 10\,\mu$m and $v_e = 0.02$, and an optical thickness of 10. Reflectances values are plotted as a function of the scattering angle, which allow us to obtain the characteristic angular signature of water clouds with the cloudbow near 140° and all the supernumary bows observable between 145° and 165°. Error bars corresponding to the statistical errors of the Monte–Carlo simulations are also plotted, which allows us to be sure that the deviation observed between 3-D and 1-D are real and outside the statistical noise of the computation. In the case of the polarized reflectances, we computed them from ΔQ and ΔU using:

$$\Delta I_p = \frac{|Q|\Delta Q + |U|\Delta U}{I_p} \,. \tag{3.27}$$

The non-linearity of reflectances as a function of the optical thickness implies the so-called plane-parallel bias (Fig. 3.7) (Cahalan et al., 1994), which means

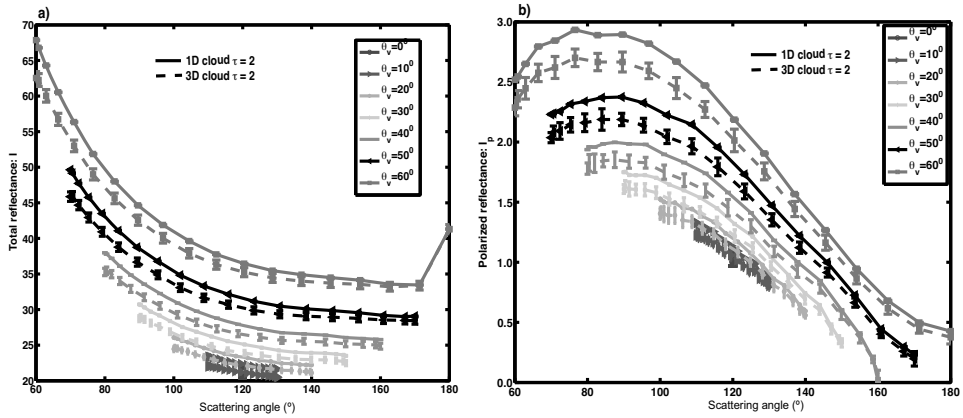


Fig. 3.10. Total (a) and polarized (b) reflectances for $\theta_s = 60°$ as a function of the scattering angles ($\Theta = 180°$ corresponding to backward scattering). Solid lines are for a homogeneous cirrus cloud of optical thickness 2 and dashed lines for the spatial mean reflectance values for the cirrus cloud presented in Fig. 3.8. Error bars corresponds to the statistical error of Monte–Carlo computations. Modified from Cornet et al. (2009). [For the color version of this figure, see the color section.]

that the 3-D mean total reflectances are smaller than the corresponding 1-D homogeneous reflectances in the two cases (Fig. 3.10(a) and Fig. 3.11(a)) for all the scattering angles. Concerning the cirrus cloud, the relative differences for total reflectances are between 0 and −8% and tend to decrease with the scattering angle. In Fig. 3.10(b), we note that this overestimation of reflectances under the 1D assumption exists also for the polarized reflectances of the cirrus cloud (Fig. 3.10(b)) with relative differences ranging from 7% to 8% whatever the viewing direction. Concerning the water cloud, the relative differences range between −2% and −14% for the total reflectances and in contrast to the cirrus cloud, overestimation of polarized reflectances under the homogeneous assumption does not exist (Fig. 3.11(b)). However, we can notice that the statistical error is important and does not allow us to be sure of the result. This is due to the strong forward peak of the Mie phase function which leads to very noisy reflectances fields. This can be avoided using a truncated phase function. Studies need therefore to be pursued to confirm this result and to understand why the plane-parallel bias appears only for the cirrus case. But, as polarized reflectances saturate quite quickly with optical thickness, this may be due to the difference in optical thickness values which are larger for the stratocumulus cloud.

Nevertheless, this result could be very interesting for cloud water microphysical retrieval because it tends to show that the sub-pixel optical thickness heterogeneity does not modify the microphysical signature of liquid water cloud. The retrieval of cloud size particles based on the angular signature (Bréon and Goloub, 198; Bréon and Doutriaux-Boucher, 2005) would thus give correct results even in the presence of variable cloud, as long as they have sufficient optical thickness. However, as macrophysical heterogeneity is often linked with microphysical heterogeneity, other studies are necessary to get a complete understanding of the heterogeneity effect on the polarized signal.

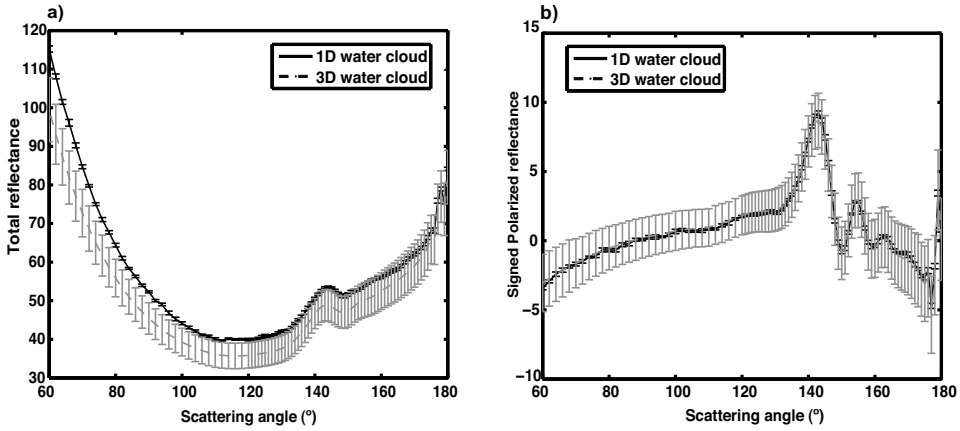


Fig. 3.11. Comparisons between 1-D and 3-D computations of total (a) and polarized reflectances (b) as a function of the scattering angles for the stratocumulus cloud presented in Fig. 3.8(c) and (d). Error bars correspond to the statistical error of Monte–Carlo computations.

3.5 Conclusions and perspectives

As polarized measurements are very suitable for obtaining properties of atmospheric components and as several instruments are currently developed to make these type of measurements, it becomes very important to have realistic atmospheric radiative transfer model handling the polarization and able to reproduce all the complexity of the atmosphere and especially of cloudy atmosphere. In Cornet et al. (2009), we presented a model based on Monte–Carlo methods, which allows us to simulate total and polarized reflectances above tri-dimensional medium. In this chapter, we recall the general scheme of this Monte–Carlo model and especially the particularities introduced by the account of the polarization. This model is tested without polarization for 3-D cases by comparing results with SHDOM simulations and is successfully validated with polarization for homogeneous cloud cases by comparing it with an Adding-Doubling model. In order to have a complete model, it remains to include, in the near future, the surface polarization in the calculation.

In this chapter, examples of 3-D radiative transfer inside clouds are presented. The clouds were created with a stochastic cloud model (Szczap, 2009). Total and polarized reflectances of two clouds, a cirrus cloud and a water cloud, were computed and analyzed. The results show that polarized reflectances can be sensitive to 3-D effects in the same way as total reflectances.

For the cirrus cloud, at high resolution and with an oblique sun illumination, we observed 3-D effects such as illumination and shadowing effects which respectively enhance or reduce the reflectances. Values higher than those predicted by the homogeneous assumption can thus be reached. In addition, at a medium resolution of 10 km, total as well as polarized reflectances of the cirrus cloud are overestimated under the homogeneous assumption because of the so-called plane-parallel bias. Concerning the water cloud with a mean optical thickness of 10, in contrast, and surprisingly, the plane-parallel bias does not appear for polarized reflectances. This may be due to the larger optical thicknesses used, given that the polarized reflectances saturate for quite small values of it. The average polarized reflectance is thus equal to the value obtained with the homogeneous model. This can be an important result for issues concerning cloud size particles retrieval because, if it is confirmed, that means that, given a sufficient resolution, microphysics retrieval from multi-viewing measurements might not be influenced by the sub-pixel optical thickness heterogeneity.

In this chapter, only two examples were studied; the work needs to be pursued with other kind of clouds, including fractional cloud cover and clouds with microphysical heterogeneity. The model used in this chapter is an essential tool to get a better knowledge of the limitation of using the 1-D assumption to infer cloud microphysics from multi-viewing polarized measurements and to exploit measurements of polarized reflectances that will be acquired by the future sensors.

Acknowledgments

This research has been supported by the CNES, the CNRS, the University of Lille 1 and the French ‘Programme National de Télédétection Saptiale’ (PNTS). The

authors thank Dr Alexander Kokhanovsky and the anonymous reviewer who read the chapter very carefully and made very constructive remarks and comments.

References

Auriol, F., Leon, J.-F., Balois, J.-Y., Verwaerde, C., François, P., Riedi, J., Parol, F., Waquet, F., Tanré, D., and Goloub, P. (2008). Multidirectional visible and shortwave infrared polarimeter for atmospheric aerosol and cloud observation: OSIRIS (Observing System Including PolaRisation in the Solar Infrared Spectrum). In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 7149. Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference.

Avery, L. W. and House, L. L. (1969). The monte carlo technique applied to radiative transfer. *Journal of Quantitative Spectroscopie and Radiative Transfer*, **9**:1579–1591.

Bréon, F.-M. and Doutriaux-Boucher, M. (2005). A comparison of cloud droplet radii measured from space. *IEEE Transactions on Geoscience and Remote Sensing*, **43**:1796–1805.

Bréon, F.-M. and Goloub, P. (1998). Cloud droplet effective radius from spaceborne polarization measurements. *Geophysical Research Letters*, **25**:1879–1882.

C-Labonnote, L., Brogniez, G., Buriez, J.-C., Doutriaux-Boucher, M., Gayet, J.-F., and Macke, A. (2001). Polarized light scattering by inhomogeneous hexagonal monocrystals. Validation with ADEOS-POLDER measurements. *Journal of Geophysical Research (Atmospheres)*, **106**:12139–12153.

Cahalan, R. F., Ridgway, W., Wiscombe, W. J., Bell, T. L., and Snider, J. B. (1994). The albedo of fractal stratocumulus clouds. *Journal of Atmospheric Sciences*, **51**:2434–2460.

Cashwell, E. D. and Everett, C. J. (1959). *A Practical Manual on the Monte-Carlo Method for Random Walk Problems*. Pergamon, Oxford.

Chandrasekar, S. (1960). *Radiative Transfer*. Dover, New York.

Chepfer, H., Goloub, P., Riedi, J., De Haan, J. F., Hovenier, J. W., and Flamant, P. H. (2001). Ice crystal shapes in cirrus clouds derived from POLDER/ADEOS-1. *Journal of Geophysical Research (Atmospheres)*, **106**:7955–7966.

Cornet, C. and Davies, R. (2008). Use of MISR measurements to study the radiative transfer of an isolated convective cloud: Implications for cloud optical thickness retrieval. *Journal of Geophysical Research (Atmospheres)*, **113**:D04202.

Cornet, C., C-Labonnote, L., and Szczap, F. (2010). Three-dimensional polarized Monte Carlo atmospheric radiative transfer model (3DMCPOL): 3D effects on polarized visible reflectances of a cirrus cloud. *Journal of Quantitative Spectroscopie and Radiative Transfer*, **111**, doi: 10.1016/j.jqsrt.2009.06.013.

de Haan, J. F., Bosma, P. B., and Hovenier, J. W. (1987). The adding method for multiple scattering calculations of polarized light. *Astronomy and Astrophysics*, **183**:371–391.

Deschamps, P.-Y., Breon, F.-M., Leroy, M., Podaire, A., Bricaud, A., Buriez, J.-C., and Seze, G. (1994). The POLDER mission: instrument characteristics and scientific objectives. *IEEE Transactions on Geoscience and Remote Sensing*, **32**:598–615.

Emde, C., Buras, R., Mayer, B., and Blumthaler, M. (2009). The impact of aerosols on polarized sky radiance: Model development, validation, and applications. *Atmospheric Chemistry and Physics*, submitted.

Evans, K. F. (1998). The spherical harmonics discrete ordinate method for three-dimensional atmospheric radiative transfer. *Journal of Atmospheric Sciences*, **55**:429–446.

Evans, K. F. and Marshak, A. (2005). Numerical methods, Ch. 4 in *3D Radiative Transfer in Cloud Atmospheres*. Springer-Verlag, New York.

Ferlay, N. and Isaka, H. (2006). Multiresolution analysis of radiative transfer through inhomogeneous media. Part I: Theoretical development. *Journal of Atmospheric Sciences*, **63**:1200–1212.

Goloub, P., Herman, M., Chepfer, H., Riedi, J., Brogniez, G., Couvert, P., and Séze, G. (2000). Cloud thermodynamical phase classification from the POLDER spaceborne instrument. *Journal of Geophysical Research (Atmospheres)*, **105**:14747–14760.

Hansen, J. E. and Travis, L. D. (1974). Light Scattering in Planetary Atmospheres, *Space Science Review*, **16**:527–610.

Herman, M., Deuze, J.-L., Marchand, A., Roger, B., and Lallart, P. (2005). Aerosol remote sensing from POLDER/ADEOS over the ocean: Improved retrieval using a nonspherical particle model. *Journal of Geophysical Research (Atmospheres)*, **110**:D10502.

Hovenier, J. W. (1969). Symmetry relationships for scattering of polarized light in a slab of randomly oriented particles. *Journal of Atmospheric Sciences*, **26**:488–499.

Iwabuchi, H. (2006). Efficient Monte Carlo methods for radiative transfer modeling. *Journal of Atmospheric Sciences*, **63**:2324–2339.

Iwabuchi, H. and Hayasaka, T. (2002). Effects of cloud horizontal inhomogeneity on the optical thickness retrieved from moderate-resolution satellite data. *Journal of Atmospheric Sciences*, **59**:2227–2242.

Lenoble, J., Herman, M., Deuzé, J. L., Lafrance, B., Santer, R., and Tanré, D. (2007). A successive order of scattering code for solving the vector equation of transfer in the earth's atmosphere with aerosols. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **107**:479–507.

Loeb, N. G. and Coakley, Jr., J. A. (1998). Inference of marine stratus cloud optical depths from satellite measurements: Does 1D theory apply? *Journal of Climate*, **11**:215–233.

Marchuk, G. I., Mikhailov, G. A., and Nazaraliev, M. A. (1980). *The Monte Carlo Methods in Atmospheric Optics*. Springer, New York.

Marshak, A. and Davis, A. (2005). *3D Radiative Transfer in Cloudy Atmospheres*. Springer, New York.

Marshak, A., Platnick, S., Varnai, T., Wen, G., and Cahalan, R. F. (2006). Impact of three-dimensional radiative effects on satellite retrievals of cloud droplet sizes. *Journal of Geophysical Research (Atmospheres)*, **111**:D09207.

Mayer, B. (2009). Radiative transfer in the cloudy atmosphere. *European Physical Journal Conferences*, **1**:75–99.

Mishchenko, M. I., Cairns, B., Kopp, G., Schueler, C. F., Fafaul, B. A., Hansen, J. E., Hooker, R. J., Itchkawich, T., Maring, H. B., and Travis, L. D. (2007). Accurate monitoring of terrestrial aerosols and total solar irradiance: Introducing the Glory mission. *Bulletin of the American Meteorological Society*, **88**:677-691.

Pincus, R. and Evans, K. (2009). Computational cost and accuracy in calculating three-dimensional radiative transfer: Results for new implementations of Monte Carlo and SHDOM. *Journal of Atmospheric Sciences*, **66**:331-3146.

Ramella-Roman, J. C., Prahl, S. A., and Jacques, S. L. (2005). Three Monte Carlo programs of polarized light transport into scattering media: Part I. *Optics Express*, **13**:10392–10405.

Riedi, J., Goloub, P., and Marchand, R. T. (2001). Comparison of POLDER cloud phase retrievals to active remote sensors measurements at the ARM SGP site. *Geophysical Research Letters*, **28**:2185–2188.

Rozanov, V. V. and Kokhanovsky, A. A. (2006). The solution of the vector radiative transfer equation using the discrete ordinates technique: Selected applications. *Atmospheric Research*, **79**:241–265.

Stamnes, K., Tsay, S.-C., Jayaweera, K., and Wiscombe, W. (1988). Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media. *Applied Optics*, **27**:2502–2509.

Sun, W., Loeb, N. G., and Yang, P. (2006). On the retrieval of ice cloud particle shapes from POLDER measurements. *Journal of Quantitative Spectroscopy and Radiative Transfer*, **101**:435–447.

Szczap, F. (2009). 3DCloud, a fast and flexible generator of optical properties of cirrus and stratocumulus/small cumulus based on a simplified dynamical, thermodynamic and stochastic framework. *Atmospheric Chemistry and Physics*, submitted.

Szczap, F., Isaka, H., Saute, M., Guillemet, B., and Iotukhovski, A. (2000). Effective radiative properties of bounded cascade absorbing clouds: Definition of an effective single-scattering albedo. *Journal of Geophysical Research (Atmospheres)*, **105**:20635–20648.

Van de Hulst, H. C. (1980). *Multiple Light Scattering. Tables, Formulas and Applications*. Academic, New York.

Varnai, T. and Davies, R. (1999). Effects of cloud heterogeneities on shortwave radiation: Comparison of cloud-top variability and internal heterogeneity. *Journal of Atmospheric Sciences*, **56**:4206–4224.

Varnai, T. and Marshak, A. (2001). Statistical analysis of the uncertainties in cloud optical depth retrievals caused by three-dimensional radiative effects. *Journal of Atmospheric Sciences*, **58**:1540–1548.

Varnai, T. and Marshak, A. (2002). Observations of three-dimensional radiative effects that influence MODIS cloud optical thickness retrievals. *Journal of Atmospheric Sciences*, **59**:1607–1618.

Zinner, T. and Mayer, B. (2006). Remote sensing of stratocumulus clouds: Uncertainties and biases due to inhomogeneity. *Journal of Geophysical Research (Atmospheres)*, **111**.

4 Linearization of radiative transfer in spherical geometry: an application of the forward-adjoint perturbation theory

Holger H. Walter and Jochen Landgraf

The remote sensing of atmospheric constituents with limb-viewing satellite instruments or with nadir viewing instruments at large solar zenith angles requires a forward model that simulates the backscattered radiance taking the spherical shape of the Earth atmosphere into account. In addition, many retrieval schemes are based on a linearization of such a forward model. Whenever it is important to take multiple scattering into account (e.g. due to light scattering air molecules, aerosols and clouds) the linearization of the measurement simulation with respect to the parameters to be retrieved is not trivial. Here, the forward-adjoint perturbation theory provides a general method to linearize radiative transfer. In the first part of this review chapter we provide the theoretical background of the linearization approach for a radiative transfer problem in a spherical model atmosphere which is illuminated by a collimated solar beam. Using an operator formulation of radiative transfer allows one to express the linearization approach in a universally valid notation. Depending on the particular formulation of the radiative transfer problem the perturbation of internal sources has to be taken into account in addition. The needed adjoint calculation corresponds to a so-called searchlight problem that requires the use of three-dimensional radiative transfer simulations in general. Subsequently we show how symmetries of the forward radiation field and a proper choice of the radiation sources can be used to simplify the needed adjoint calculations substantially.

As an example we present the linearization of a numerical radiative transfer model for a spherical shell atmosphere in the second part of this article. It allows the interpretation of limb measurements in the ultraviolet and visible spectral range. Here, the scalar radiative transfer problem is solved in its forward and adjoint formulation. Subsequently, we apply the perturbation theory approach in the calculation of the derivatives of the radiance at the top of the atmosphere with respect to the absorption properties of a trace gas species in the case of a limb-viewing satellite instrument.

4.1 Introduction

Current and future satellite instruments, which observe the Earth's atmosphere in limb viewing geometry in the ultraviolet (UV) and visible (VIS) spectral range,

provide information about the distribution of atmospheric trace constituents with high vertical resolution. For example the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) [5], launched in 2002 on board of ESA's ENVISAT satellite, measures backscattered limb radiances in the wavelength range from 240 to 2380 nm. In limb-viewing geometry, SCIAMACHY vertically scans the atmosphere from ground level up to an altitude of approximately 100 km in steps of 3 km at tangent height. This enables the retrieval of a variety of atmospheric trace constituents in high vertical resolution. Furthermore, SCIAMACHY offers the unique possibility of observing an air volume in nadir as well as in limb-viewing geometry. It is expected that the combination of both viewing geometries will facilitate a higher accuracy in the retrieval of atmospheric trace gas profiles in the troposphere and stratosphere [56]. In 2001 the Canadian Optical Spectrograph and Infrared Imager System (OSIRIS) [38] was launched on board the Odin satellite. The instrument observes the backscattered radiances in limb viewing geometry and comprises a UV/VIS spectrograph covering the wavelength range from 280 to 800 nm and three 30 nm wide infrared (IR) channels at 1260 nm, 1270 nm and 1280 nm. The vertical distribution of e.g. O_3, NO_2, OClO, BrO and aerosols can be retrieved from OSIRIS limb measurements. Also, the Ozone Mapping and Profiler Suite (OMPS) [14], to be launched on NPOESS in 2010 by the United States of America, is dedicated to monitoring the global ozone distribution from space. This instrument consists of a nadir sensor with a wide field-of-view telescope and a limb sensor. The nadir spectrometer observes the backscattered radiance between 300 and 380 nm, whereas the limb sensor measures the along-track limb-scattered solar radiance in the spectral range from 290 to 1000 nm.

In order to retrieve an atmospheric parameter x from such limb observations with standard least squares methods [39, 41, 51], one has to simulate both the radiance I^{TOA} at the top of the model atmosphere (TOA) and its derivative with respect to the parameter to be retrieved. This corresponds to a linearization of the radiative transfer problem, given by a Taylor expansion around a first-guess parameter x_0:

$$I^{\mathrm{TOA}}(x) \;=\; I^{\mathrm{TOA}}(x_0) + \frac{\partial I^{\mathrm{TOA}}}{\partial x}(x_0)\Delta x + \mathcal{O}(\Delta x^{\,2}). \tag{4.1}$$

Here, $\mathcal{O}(\Delta x^2)$ represents higher-order terms in $\Delta x = x - x_0$.

For the simulation of limb radiance measurements in the UV and VIS spectral range, radiative transfer models, which take the multiple scattering of radiation in a spherical model atmosphere into account, are required. In general, such spherical radiative transfer models are complex and computationally highly expensive. The Monte Carlo technique [29] provides a straightforward approach to solving this type of radiative transfer problem (see e.g. [1, 11, 36, 40, 50]). However, due to its statistical nature, Monte Carlo models require a considerable computational effort in order to achieve a reasonable accuracy. Therefore alternative techniques have been developed in order to solve the spherical radiative transfer equation in a more efficient manner. [46] gives a general overview about different solution techniques in curved media. Here, applications such as radiative transfer in stellar atmospheres are discussed. For the simulation of radiative transfer in a spherical planetary atmosphere, [49] and [48] suggested an efficient solution approach,

in which the first-order scattering is calculated exactly, whereas the higher-order scattering is approximated by the use of the first two terms of a Legendre expansion of the scattering phase function only. [47] applied this method to determine the source function and brightness of a spherical planetary atmosphere. Furthermore, several radiative transfer approaches have been suggested [2, 3, 12, 24] for the simulation of the azimuthally-averaged radiation field in a spherical atmosphere. However, remote sensing in limb-viewing geometry requires the simulation of the azimuthal dependence of the radiation field. This angular dependence is fully taken into account by the model of [18], which employs a Gauss–Seidel iteration scheme to solve the integral form of the radiative transfer equation. Here, the assumption of a constant ratio of single scattering to multiple scattering along a conical boundary facilitates the efficient calculation of the multiple-scattering source function in spherical geometry. However, due to this approximation the radiation field can only be calculated in cases in which the single-scattering terms are not zero. [42] presented an alternative radiative transfer model which is based on a Picard iteration scheme. A radiation field, which is calculated in pseudo-spherical approximation, serves as an initial guess in order to integrate the radiative transfer equation along discrete lines in spherical geometry in subsequent iterations. Corresponding methods have been developed independently in order to solve the multiple scattering problem in spherical geometry [13, 31, 59]. [27] shows a quantitative comparison of currently available radiative transfer models, suited for the simulation of limb radiance measurements.

Minor attention has been given to the linearization of spherical radiative transfer. The linearization of the spherical radiative transfer models has been performed either by the finite difference method or by approximate methods so far. In the finite difference method (see e.g. [31]) two radiative transfer calculations are utilized to approximate the derivative of I^{TOA}: one for an unperturbed atmosphere and one for an atmosphere for which the desired atmospheric parameter x has been perturbed by a small amount. For trace gas profile retrieval such an approach leads to many repetitive computations, as for every perturbation of the trace gas concentration at a different altitude a new radiative transfer calculation is required. The combination of this method with the complex spherical radiative transfer models in an iterative retrieval approach requires considerable computational resources. Therefore the use of a (quasi-)analytical linearization of spherical radiative transfer is desirable. The pseudo-spherical approximation, for example, found a wide application in the remote sensing of atmospheric parameters (compare e.g. [43, 44]). Furthermore, [19] has presented a linearized spherical radiative transfer model, taking two orders of scattering into account, which is suited for a concise trace gas profile retrieval in atmospheric scenarios with a low fraction of multiple scattered radiation. Recently, [40] presented a linearized vector radiative transfer model for a spherically symmetrical atmosphere, which is based on the Monte Carlo estimation technique [29]. Therein the derivative of the radiance at TOA with respect to a trace gas number density is calculated using the same Markov chain as for the calculation of the radiance itself, which makes the model very efficient. Details may be found in [29] and [40].

Alternatively the linearization of a radiative transfer problem can be achieved using the forward-adjoint perturbation theory, which is known from neutron trans-

port theory [4]. This technique can be used for the linearization with respect to absorption and scattering properties of the atmosphere. In addition, the approach is independent of the solution method of the radiative transfer equation and therefore any radiative transfer model, which simulates the internal forward and adjoint radiation field, can be linearized by means of the method presented here. Moreover, the perturbation theory can be applied in a straightforward manner to radiative transfer problems, which take the polarization of radiation into account. The perturbation theory approach was firstly used by [28], who applied it to plane-parallel atmospheric radiative transfer. A great variety of different applications of the forward-adjoint perturbation theory in plane-parallel geometry has been developed (e.g. [6], [53–55], [45], [23], [17] and [16]). Due to the fact that the perturbation theory approach can be formulated for an arbitrary geometry [4], it can be applied also to the full spherical radiative transfer problem in a straightforward manner. This was demonstrated by [57] and [59]. In this review article we give a detailed outline how to employ the forward-adjoint perturbation theory in spherical geometry.

In Section 4.2 the forward and the adjoint formulations of radiative transfer in spherical geometry are introduced. Both formulations of radiative transfer are needed in subsection 4.2.3 to derive an analytical linearization of the spherical radiative transfer problem. Any spherical radiative transfer model, which simulates the internal forward and adjoint radiation field, can be linearized by means of the approach presented here. Section 4.3 discusses symmetries which can be used to simplify the calculation of the forward and adjoint radiation field. Finally, in Section 4.4 the linearization with respect to the amount of trace gases at different altitudes is demonstrated for a simplified spherical radiative transfer problem, which yields a radiation field that does not depend on the global azimuth.

4.2 Forward-adjoint perturbation theory in spherical geometry

4.2.1 The forward radiative transfer equation

We consider a planetary atmosphere that is illuminated symmetrically by parallel solar beams in downward direction of the global zenith Z_S (see Fig. 4.1). As we are interested in the ultraviolet and visible part of the solar spectrum, thermal emission will not be taken into account in the following description of the radiative transfer processes. Any refractional effects of radiation are neglected, therefore it can be assumed that light travels along straight lines. Because a scalar approximation of radiative transfer simplifies the notation of the associated equations substantially, polarization is not taken into account in the remainder. Depending on atmospheric composition, observation geometry and wavelength, this approximation, however, can lead to errors of up to 10% in the simulated radiance [32]. For these cases the scalar radiative transfer equation and also the corresponding expressions for the forward-adjoint perturbation theory can be extended to a vector formulation in a straightforward manner, as is shown by [15].

Basically, the radiative transfer equation is a detailed balance equation which locally describes the change of radiation in a certain direction due to sinks and

sources in the atmosphere. The loss of radiation is caused by extinction processes as a result of absorption and scattering of photons out of the propagation direction of the light. Sources of radiation are the scattering of photons from other directions into the propagation direction, emission sources within the atmosphere and external radiation sources like the sun. The time-independent, monochromatic radiative transfer equation for an isotropic medium in local thermodynamic equilibrium can be written in operator form [6, 28] as

$$\mathbf{L}I \;=\; S_0, \tag{4.2}$$

where $I(\mathbf{r}, \mathbf{\Omega})$ is the spectral radiance at position $\mathbf{r}$ in direction $\mathbf{\Omega}$ and S_0 is the solar radiation source. The vectors $\mathbf{r}$ and $\mathbf{\Omega}$ constitute the phase space of the radiative transfer problem. For radiative transfer in spherical geometry it is convenient to represent the position $\mathbf{r}$ in the atmosphere by spherical coordinates, i.e. the radius r, the global zenith angle Ψ and the global azimuth angle Φ. The transport operator $\mathbf{L}$ is given by

$$\mathbf{L} \;=\; \int\limits_{4\pi} d\tilde{\Omega}\,\{\,\left[\frac{d}{ds} + \beta_e(\mathbf{r})\right]\delta(\mathbf{\Omega} - \tilde{\mathbf{\Omega}}) - \frac{\beta_s(\mathbf{r})}{4\pi} P(\mathbf{r}, \tilde{\mathbf{\Omega}}, \mathbf{\Omega})\,\}\,, \tag{4.3}$$

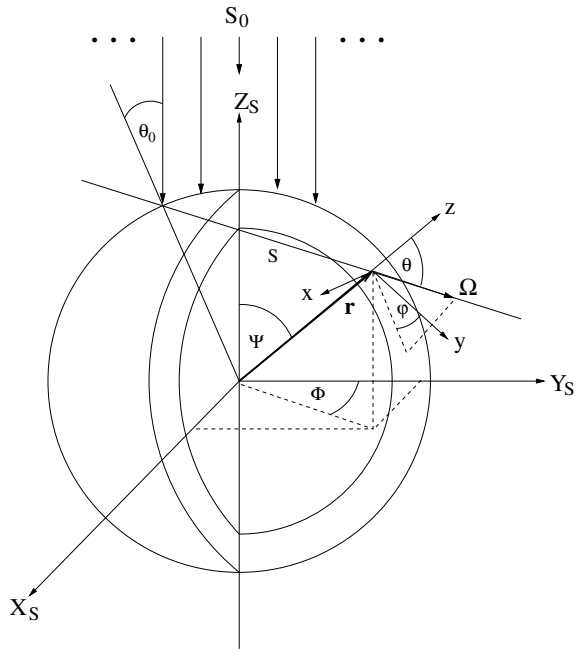


Fig. 4.1. Schematic overview of the geometry used for the description of radiative transfer in a spherical atmosphere. In a spherical coordinate system the radiation field is a function of the three spatial coordinates r, Ψ, Φ and the two directional coordinates θ and φ. The incident solar radiation illuminates the upper hemisphere of the planet. Here, the solar zenith angle θ_0 is defined locally at the top of the model atmosphere with respect to the inner normal.

where d/ds is the derivative along the pathlength s in direction $\mathbf{\Omega}$ and may be also expressed as $\mathbf{\Omega} \cdot \nabla$. The directional vector $\mathbf{\Omega} = (\theta, \varphi)$ is specified by the local zenith angle θ, defined with respect to the outer normal, and the local azimuth angle φ. Further $d\tilde{\Omega} = \sin\tilde{\theta} d\tilde{\theta}\, d\tilde{\varphi}$. δ is the Dirac-delta function with

$$\delta(\mathbf{\Omega} - \tilde{\mathbf{\Omega}}) \;=\; \frac{1}{\sin\theta}\delta(\theta - \tilde{\theta})\;\delta(\varphi - \tilde{\varphi}). \tag{4.4}$$

In the transport operator (4.3) surface reflection has been neglected. Such effects, however, may be included in a straightforward manner, as shown by [22]. Finally, β_e is an extinction coefficient, β_s is a scattering coefficient and P denotes the scattering phase function. In general all three quantities vary depending on position.

The solar radiation source S_0, which illuminates the upper hemisphere of the Earth's atmosphere (see Fig. 4.1), may be expressed as

$$S_0(\mathbf{r}, \mathbf{\Omega}) \;=\; \mu_0 F_0 \delta(r - r_{\mathrm{TOA}})\delta(\mathbf{\Omega} - \mathbf{\Omega}_0)\Theta\,(\mu_0)\,. \tag{4.5}$$

Here, F_0 is the extraterrestrial flux, r_{TOA} is the distance from the center of the Earth to the top of the model atmosphere and $\mathbf{\Omega}_0 = (\theta_0, \varphi_0)$ describes the direction of the solar beam. Here, θ_0 is the local solar zenith angle and φ_0 the local solar azimuth angle. Further, θ_0 is defined with respect to the inner normal, in order to ensure comparability with common plane-parallel radiative transfer theory. μ_0 is the cosine of the local solar zenith angle θ_0. The Heaviside step function $\Theta\,(\mu_0)$ in (4.5) accomplishes the illumination of only one hemisphere of the Earth's atmosphere.

The radiation field I is subject to the free surface boundary conditions of no incoming radiation at the top (TOA) and bottom (BOA) of the atmosphere,

$$\begin{aligned} I(\mathbf{r}_{\mathrm{TOA}}, \mathbf{\Omega}) &= 0 \quad \text{for} \quad -1 \le \mu < 0 \\ I(\mathbf{r}_{\mathrm{BOA}}, \mathbf{\Omega}) &= 0 \quad \text{for} \quad 0 < \mu \le 1. \end{aligned} \tag{4.6}$$

Here, $\mu = \cos\theta$.

For a solution of (4.2) three-dimensional radiative transfer calculations are needed in general. Here, the streaming term in (4.3) can be expanded in spherical coordinates (see e.g. [42] and references therein) which yields

$$\begin{aligned} \frac{d}{ds} \;=\; & \cos\theta\frac{\partial}{\partial r} + \frac{\sin\theta\cos\varphi}{r}\frac{\partial}{\partial \Psi} + \frac{\sin\theta\sin\varphi}{r\sin\Psi}\frac{\partial}{\partial \Phi} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta} \\ & - \frac{\sin\theta\sin\varphi\cot\Psi}{r}\frac{\partial}{\partial\varphi}. \end{aligned} \tag{4.7}$$

Because of the complex structure of the streaming term, up to now no solution method other than the Monte Carlo technique [29] has been available to solve the radiative transfer problem in a three-dimensional inhomogeneous, spherical atmosphere. The Monte Carlo technique is with respect to the geometry and the atmospheric composition the most versatile approach [36, 40, 50]. Other solution methods – as discussed in the introduction – generally make use of different simplifications of the radiative transfer problem.

Once the internal radiance field $I(\mathbf{r}, \mathbf{\Omega})$ has been calculated with the help of a numerical radiative transfer model, any radiative effect E – in general the observable – might be obtained by an evaluation of the inner product [6, 28]

$$E \;=\; \langle R|I\rangle, \tag{4.8}$$

where the inner product describes an integration over the full phase space [4], viz.

$$\langle R|I\rangle \;=\; \int_V dV \int_{4\pi} d\Omega \;\; R(\mathbf{r},\mathbf{\Omega})I(\mathbf{r},\mathbf{\Omega}). \tag{4.9}$$

Here, dV represents a differential volume element. R is the so-called response function which is defined by the measurement. In the context of satellite remote sensing, the radiative effect under investigation is the radiance at the top of the model atmosphere in the viewing direction of the instrument, thus $E = I(\mathbf{r}_\mathrm{v},\mathbf{\Omega}_\mathrm{v}) \equiv I^\mathrm{TOA}$. Here, $\mathbf{\Omega}_\mathrm{v} = (\theta_\mathrm{v},\varphi_\mathrm{v})$ describes the viewing direction of the satellite instrument, where θ_v is the viewing zenith angle and φ_v is the viewing azimuth angle, both defined at the location of the satellite at $\mathbf{r}_\mathrm{v}$.[1] For a measurement of I^TOA with a satellite, the appropriate response function is given by the direction of the instrument's line-of-sight

$$R(\mathbf{r},\mathbf{\Omega}) \;=\; \delta(\mathbf{r}-\mathbf{r}_\mathrm{v})\delta(\mathbf{\Omega}-\mathbf{\Omega}_\mathrm{v}), \tag{4.10}$$

with

$$\delta(\mathbf{r}-\mathbf{r}_\mathrm{v}) = \delta(r-r_\mathrm{TOA})\,\frac{1}{r}\delta(\Psi-\Psi_\mathrm{v})\,\frac{1}{r\sin\Psi}\delta(\Phi-\Phi_\mathrm{v}) \tag{4.11}$$

in the representation of spherical coordinates.

4.2.2 The adjoint formulation of radiative transfer

The linearization of spherical radiative transfer with the forward-adjoint perturbation theory requires the solution of an adjoint transfer equation in addition. The adjoint transfer equation corresponds to a backward formulation of radiative transfer in space and can be derived from the forward radiative transfer equation (4.2). In order to derive the adjoint transfer equation, we refer to the theoretical approach introduced by [4]. (See also [57] for details.)

The adjoint transfer equation needed for a linearization is given in operator notation by

$$\mathbf{L}^\dagger I^\dagger \;=\; R, \tag{4.12}$$

where $I^\dagger$ denotes the adjoint intensity field, R is the response function from (4.10) and $\mathbf{L}^\dagger$ is the transport operator adjoint to the forward transport operator $\mathbf{L}$. The adjoint transport operator $\mathbf{L}^\dagger$ is defined by the following equation:

$$\langle \mathbf{L}^\dagger I_2|I_1\rangle \;=\; \langle I_2|\mathbf{L}I_1\rangle \tag{4.13}$$

for two arbitrary intensity fields I_1 and I_2. It can be shown that the adjoint operator is given by [4]

[1] Here, $\mathbf{r}_\mathrm{v}$ is given by the radius r_TOA, the global zenith angle Ψ_v and the global azimuth angle Φ_v.

$$\mathbf{L}^{\dagger} \;=\; \int\limits_{4\pi} d\tilde{\Omega}\,\{\,\left[-\frac{d}{ds}+\beta_e(\mathbf{r})\right]\delta(\mathbf{\Omega}-\tilde{\mathbf{\Omega}})-\frac{\beta_s(\mathbf{r})}{4\pi}P(\mathbf{r},\mathbf{\Omega},\tilde{\mathbf{\Omega}})\,\}\,. \tag{4.14}$$

Thus, in the adjoint problem the sign of the streaming term d/ds is negative and the directions in the scattering phase function are reversed. Like the forward transport operator, the adjoint transport operator includes all scattering and extinction processes of the adjoint problem. However, the adjoint operator $\mathbf{L}^{\dagger}$ also differs from the forward transport operator $\mathbf{L}$; in other words $\mathbf{L}$ is not self-adjoint.

The adjoint intensity $I^{\dagger}$ has to satisfy the boundary conditions of no outgoing adjoint intensity at the top and the bottom of the atmosphere,

$$\begin{aligned} I^{\dagger}(\mathbf{r}_{\mathrm{TOA}},\mathbf{\Omega}) &= 0 \quad \text{for} \quad 0 \le \mu < 1 \\ I^{\dagger}(\mathbf{r}_{\mathrm{BOA}},\mathbf{\Omega}) &= 0 \quad \text{for} \quad -1 < \mu \le 0. \end{aligned} \tag{4.15}$$

With the help of the resulting adjoint intensity field from (4.12), the radiative effect E as defined in (4.8) can be expressed in an alternative way [6, 28]:

$$E = \langle I^{\dagger} | S_0 \rangle. \tag{4.16}$$

This shows that if the satellite response function R acts as an adjoint radiation source, the resulting adjoint intensity field has a clear physical meaning. Then, $I^{\dagger}$ can be interpreted as the importance [4, 26] of scattered photons within the atmosphere for a given measurement. In that way the adjoint field concisely describes the domain of influence of a satellite observation. Hence, the adjoint problem and the forward problem become closely related.

In general, the solution of (4.12) also requires three-dimensional radiative transfer calculations. Here, an algorithm might be needed which is independent of the algorithm to solve the forward radiative transfer problem (4.2). However, due to the reciprocal nature of light, the adjoint transfer problem can be transformed into a pseudo-forward problem, which allows one to utilize the same radiative transfer model for the solution of the forward radiative transfer equation (4.2) as well as for the solution of the adjoint transfer equation (4.12) [4, 10]. By exchanging and reversing the incoming and outgoing direction, the adjoint transport operator transforms into the forward one. That is, the lightpath can simply be reversed. This reciprocity principle can be expressed by a symmetry relation for the phase function [4], viz.

$$P(\mathbf{r},\mathbf{\Omega}_1,\mathbf{\Omega}_2) \;=\; P(\mathbf{r},-\mathbf{\Omega}_2,-\mathbf{\Omega}_1). \tag{4.17}$$

All atmospheric scatterers (i.e. molecules, aerosols, liquid water droplets and ice crystals) fulfill the symmetry relation (4.17).[2] Therefore we can substitute the adjoint intensity field $I^{\dagger}$ by the pseudo-forward intensity field ξ, which is defined as

$$\xi(\mathbf{r},\mathbf{\Omega}) \;=\; I^{\dagger}(\mathbf{r},-\mathbf{\Omega}). \tag{4.18}$$

[2] In cases where the polarization of radiation plays an important role, e.g. for the description of scattering by ice crystals, the symmetry relation (4.17) can be extended to a corresponding symmetry relation for the phase matrix. For details, see e.g. [33].

Then it is possible to write for (4.12) in its pseudo-forward formulation

$$\mathbf{L}\xi \;=\; S_\xi \tag{4.19}$$

with $\mathbf{L}$ the forward transport operator and

$$S_\xi(\mathbf{r}, \mathbf{\Omega}) = R(\mathbf{r}, -\mathbf{\Omega}) \tag{4.20}$$

as a new radiation source. Subsequently, the pseudo-forward transport equation (4.19) can be solved with the same radiative transfer model which is employed in the solution of the forward transport equation. For it only the radiation source has to be exchanged, viz. S_ξ instead of S.

The pseudo-forward formulation (4.19) is also known as the backward formulation of radiative transfer. Here, the satellite instrument is assumed to emit photons in the viewing direction into the atmosphere. A narrow light beam, which illuminates the atmosphere, could serve as an example. As a consequence, the pseudo-forward formulation defines a so-called searchlight problem. In the case of three-dimensional radiative transfer this results in an adjoint field, which is highly peaked and concentrated around the satellite's line-of-sight (see Fig. 4.2, upper right). The actual shape of this radiation field crucially depends on the absorption and scattering properties of the atmosphere.

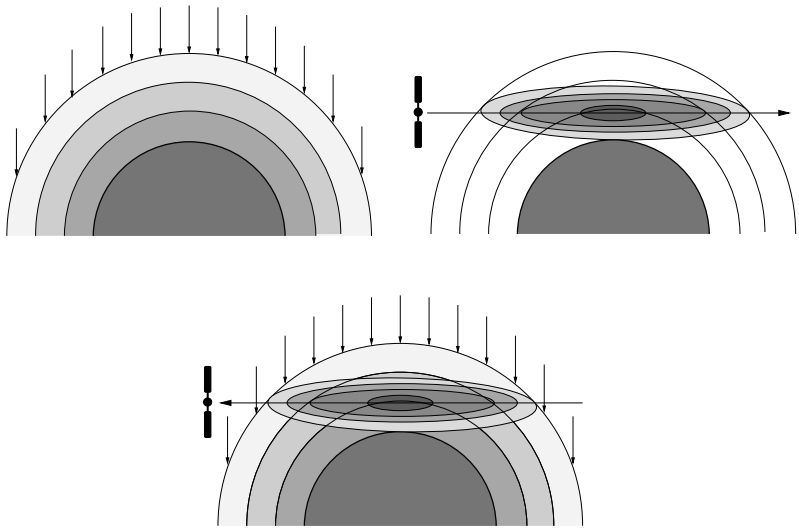


Fig. 4.2. Qualitative illustration of the forward radiation field and its importance. The left upper graph illustrates the diffuse forward radiation field due to the solar illumination at the top of the atmosphere. The graph in the upper right shows the two-dimensional section of the pseudo-forward radiation field for a limb-viewing satellite instrument. The combination of both fields in the lower graph shows where the highest sensitivity with respect to atmospheric changes exists. Here, the pseudo-forward field has been transformed in its corresponding adjoint field by a reversal of directions. Additional perturbations of internal sources and changes in the response function are not shown here, but can be illustrated in the same manner.

For further details we refer to [4] and [57]. The combination of the forward field and the adjoint field enables the calculation of a sensitivity of the measurement with respect to a perturbation of the atmosphere. Such an approach can be used to linearize any spherical radiative transfer model that calculates the internal forward and adjoint intensity fields.

4.2.3 Perturbation theory in spherical coordinates

With the solution I_0 and $I_0^\dagger$ of the forward and the adjoint transfer problem for a given atmospheric state x_0, it is possible to determine the impact of a change $\Delta x = x - x_0$ on the radiative effect E, (4.8), to first order[3] by

$$E(x) \;=\; E(x_0) + \Delta E + \mathcal{O}(\Delta x^2) \tag{4.21}$$

with

$$\Delta E \;=\; -\langle I_0^\dagger | \Delta \mathbf{L} I_0 \rangle + \langle I_0^\dagger | \Delta S \rangle + \langle \Delta R | I_0 \rangle. \tag{4.22}$$

Here, $\mathcal{O}(\Delta x^2)$ represents second- and higher-order terms. A perturbation of x_0 by an amount Δx to $x = x_0 + \Delta x$ causes a change of the transport operator $\mathbf{L}$, of the source function S and of the response funtion R by $\Delta\mathbf{L}$, ΔS, and ΔR, respectively. The derivation of (4.22) is analogous to that presented by [28], [6, 7] and [53–55].

The perturbation $\Delta\mathbf{L}$ may be calculated to first order by

$$\Delta \mathbf{L} = \mathbf{L}_0' \Delta x \tag{4.23}$$

where $\mathbf{L}_0'$ is the first derivative of $\mathbf{L}$ with respect to the state vector $\mathbf{x}$. Analogous expressions hold for the perturbation ΔS and ΔR with derivatives S_0' and R_0'. This leads to the first order approximation

$$\Delta E \;=\; -\Delta x \langle I_0^\dagger | \mathbf{L}_0' I_0 \rangle + \Delta x \langle I_0^\dagger | S_0' \rangle + \Delta x \langle R_0' | I_0 \rangle. \tag{4.24}$$

Finally, a comparison of (4.24) with a corresponding Taylor expansion of the radiative effect E around $\mathbf{x}_0$ yields the analytical expressions for the derivatives of E. Taking into account that the desired radiative effect is the radiance at the top of the atmosphere in the viewing direction of the satellite instrument, the corresponding derivatives are given by

$$\frac{\partial I^{\mathrm{TOA}}}{\partial x}(\mathbf{x}_0) \;=\; -\langle I_0^\dagger | \mathbf{L}_0' I_0 \rangle + \langle I_0^\dagger | S_0' \rangle + \langle R_0' | I_0 \rangle. \tag{4.25}$$

If the state vector $\mathbf{x}$ represents the number density of an absorbing trace gas in an atmospheric volume ΔV the derivative $\mathbf{L}_0'$ is given by

$$\mathbf{L}_0' \;=\; \Gamma(\mathbf{r}) \sigma_{\mathrm{abs}}(\mathbf{r}), \tag{4.26}$$

[3] Higher-order perturbation theory, which might be needed for large perturbations or in situations where the first-order contribution disappears, generally requires the use of a Green's function formalism, which makes its application more complex. For a discussion of higher-order perturbation theory in radiation transport please see e.g. [8], [52], [21], and [60].

with

$$\Gamma(\mathbf{r}) \;=\; \begin{cases} 1 & : \quad \mathbf{r} \in \Delta V \\ 0 & : \quad \mathbf{r} \notin \Delta V \end{cases} \tag{4.27}$$

and $\sigma_{\mathrm{abs}}(\mathbf{r})$ is the absorption cross section of the considered trace gas species. Therefore the evaluation of (4.25) – in contrast to frequently employed finite difference schemes – does not depend on the perturbation Δx itself. If S and R are defined as internal sources, which is the case for example in the pseudo-spherical approximation [58], their contribution to (4.25) due to a perturbation Δx has to be taken into account. However, if S and R are defined as surface sources, they do not experience any change due to a perturbation Δx within the atmospheric volume and the last two contributions in (4.25) vanish.

For a qualitative interpretation of (4.25) one can say, that where the importance and the forward radiation field show high values (see Fig. 4.2), a change of the atmospheric state vector will have a large effect on the satellite measurement too. As a consequence the calculated sensitivity in (4.25) will be large. Where either one of the fields or both show low values, also the sensitivity with respect to changes in the trace gas number density will be small. In case of internal sources this reasoning can be adopted in straightforward manner.

Equation (4.25) allows one to efficiently calculate derivatives for different perturbations of the atmosphere. Two radiative transfer calculations are sufficient for the complete determination of the Jacobian, because the forward and the adjoint intensity field in (4.25) stay the same for all perturbations and do not need to be recalculated. The perturbations only affect the perturbed transport operator $\Delta\mathbf{L}$, the radiation source S and the response function R, and in turn the calculation of all required derivatives is reduced to the repeated evaluation of the inner products in (4.25). Compared to the forward and the adjoint radiative transfer simulations, the computational costs for the numerical evaluation of the inner product are very low. This results in an efficient linearization approach for spherical radiative transfer compared to, for example, a finite difference scheme.

The approach presented in this subsection can be used to linearize any spherical radiative transfer model that calculates the internal forward and adjoint intensity fields. It is independent of the specific solution method and contains no further approximations. Further details on the perturbation theory approach may be found in [28], [4], [6,7], [54], [23], [58], [57] and [59]. Equation (4.25) provides the linearization of a spherical radiative transfer model with respect to a trace gas density in its most general form. Depending on the formulation of the radiative transfer problem further simplifications are possible. This will be the subject of the following section.

4.3 Symmetry properties

Symmetry properties of a model atmosphere combined with symmetries of the external illumination may ease the solution of a radiative transfer problem significantly. For instance, a common assumption in spherical radiative transfer is to consider a model atmosphere which consists of homogeneous spherical shells. Furthermore, the solar source S_0 in (4.5) shows an axial symmetry with respect to the

global zenith. It is due to both symmetries that the forward radiation field I in such a case does not depend on the global azimuth angle Φ [25,48]. Another well-known example with a high degree of symmetry is the plane-parallel radiative transfer problem assuming horizontally homogeneous layers with a uniform illumination at the top of the model atmosphere. Again, it is due to the symmetry of the model atmosphere and the symmetry of the radiation source that the forward radiation field shows a dependence only on height and on the direction of the radiation. Thus, the calculation of the forward radiation field is facilitated in both cases by virtue of the imposed symmetry relations. However, these simplifications cannot be adopted in a straight-forward manner in the calculation of the adjoint or pseudoforward radiation field due to the fundamental difference between the solar source S_0 and the response function R. In general, the response function has fewer symmetries and, in the case of the response function from (4.10), the full three-dimensional searchlight problem has to be solved, even for plane-parallel radiative transfer. Nevertheless, the symmetries of a radiative transfer problem can be used in a different manner in order to simplify the formulation of the adjoint transfer problem. Here, we make use of the fact that the response function is not uniquely determined if the forward radiation field, as depicted above, is subject to symmetry properties.

For example, for a spherical shell atmosphere that is illuminated by a collimated solar beam the resulting radiation field is symmetric around the global zenith and therefore becomes independent from the global azimuth angle Φ. In such a manner the three-dimensional radiative transfer equation reduces to a two-dimensional problem in space and the spatial dependence of the radiation field can be described by the radius r and the global zenith angle Ψ only. This eases the solution of the radiative transfer equation significantly. In turn, the original response function R from (4.10) as well as its integration over the redundant variables

$$\begin{aligned} R^*(r,\Psi,\mathbf{\Omega}) &= \frac{1}{2\pi}\int d\Phi R(r,\Psi,\Phi,\mathbf{\Omega}) \\ &= \frac{1}{2\pi r^2 \sin\Psi}\delta(r-r_{\mathrm{TOA}})\delta(\Psi-\Psi_{\mathrm{v}})\delta(\mathbf{\Omega}-\mathbf{\Omega}_{\mathrm{v}}) \end{aligned} \tag{4.28}$$

describe the same radiative effect E. Thus, in contrast to (4.8), the radiative effect E might be calculated via

$$\begin{aligned} E &= \langle R^* | I \rangle \\ &= \int\limits_V dV \int\limits_{4\pi} d\Omega \;\; R^*(r,\Psi,\mathbf{\Omega}) I(r,\Psi,\mathbf{\Omega}). \end{aligned} \tag{4.29}$$

The equivalent response function R^* represents, in comparison to the response function R from (4.10), an adjoint source which illuminates the top of the model atmosphere on a circle centered around the global zenith. This, in turn, introduces symmetries into the adjoint formulation analogous to those of the forward formulation. Therefore it suffices to perform adjoint and hence pseudo-forward calculations which obey the same simplified dependencies as the forward calculation. This corresponds to a reduction in dimensionality. Thus, R^* used as a pseudo-forward source eases significantly the model calculations in the pseudo-forward formulation of radiative transfer, (4.19). Similar, in the case of plane-parallel radiative transfer, the associated pseudo-forward or searchlight problem may be replaced by a

corresponding radiative transfer problem with uniform illumination at the top of the model atmosphere. This leads to a concise formulation of the forward-adjoint perturbation theory in plane-parallel geometry [6, 23].

Generally speaking, any symmetry of the forward radiative transfer problem can be mirrored to the adjoint formulation using the concept of an equivalent response function. This concept will be of use in the following section, in order to investigate a radiative transfer problem in simplified spherical geometry.

4.4 Linearization of a radiative transfer model for a spherical shell atmosphere by the forward-adjoint perturbation theory

In this section, we develop a numerical radiative transfer model for the interpretation of limb measurements in the ultraviolet and visible spectral range. We apply the method as presented in Sections 4.2 and 4.3 to linearize a multiple-scattering radiative transfer problem for a spherical shell atmosphere [1], which is illuminated by a collimated solar beam. There is no restriction with respect to solar or viewing geometries, which makes this approach highly versatile. To solve the corresponding radiative transfer equation one can consider its integrated form, which serves as a starting point for several iterative solution approaches [13, 18, 42, 59]. In the following the numerical solution of the forward and the adjoint radiative transfer equation is achieved by a Picard iteration scheme [42]. We compare the forward simulation with Monte Carlo reference calculations for different viewing scenarios. To verify the adjoint model we compare forward and adjoint simulations using the reciprocity principle. Subsequently, we apply the perturbation theory approach to calculate the derivatives of the radiance at the top of the atmosphere with respect to the absorption properties of a trace gas species in the case of a limb-viewing satellite instrument. The calculated derivatives are verified with finite difference calculations.

As the planetary atmosphere is approximated by homogeneous spherical shells, horizontal inhomogeneities in the optical parameters of the model atmosphere – like horizontally changing trace gas concentrations, clouds or a varying surface albedo – cannot be taken into account in the simulations. An estimation of the errors in the calculation of the radiation field due to neglecting these inhomogeneities is discussed by [35]. Therein also a parameterization to correct for the effects of a horizontally varying surface albedo for spherical shell radiative transfer models has been developed. For the retrieval of stratospheric compounds using limb measurements in the ultraviolet spectral range only, the effects of a varying surface albedo are small, and radiative transfer calculations for a spherical shell atmosphere should be sufficient in general.

4.4.1 Solution of the radiative transfer equation by a Picard iteration method

4.4.1.1 The forward radiative transfer model

In the following we describe a numerical solution of the spherical radiative transfer equation based on a Picard iteration method. Here we follow the outline given by [13].

In order to solve the radiative transfer equation (4.2) including its boundary conditions (4.6) it is convenient to shift the solar source S_0 into the boundary conditions, which yields the homogeneous ordinary differential equation (ODE)

$$\mathbf{L}I \;=\; 0, \tag{4.30}$$

with

$$\begin{aligned} &I(\mathbf{r}_{\mathrm{TOA}}, \mathbf{\Omega}) = F_0\, \delta(\mathbf{\Omega} - \mathbf{\Omega}_0) && \text{for } -1 \le \mu < 0 \\ &I(\mathbf{r}_{\mathrm{BOA}}, \mathbf{\Omega}) = 0 && \text{for } 0 < \mu \le 1. \end{aligned} \tag{4.31}$$

Note that the cosine factor μ_0 in S_0 no longer occurs in the corresponding boundary condition, which can be explained by a change of the volume source S_0 into a surface source, according to the principle of general reciprocity [9]. A detailed derivation of the boundary conditions (4.31) is given in Appendix A. After splitting the intensity field I in its direct and diffuse components I_{dir} and I_{diff}, respectively, the homogeneous ODE (4.30) can be integrated formally along a characteristic line, which is given by the straight line between the point $\mathbf{r}_1$ and a second point in direction $\mathbf{\Omega}$. This provides the direct component

$$I_{\mathrm{dir}}(\mathbf{r}_1, \mathbf{\Omega}) \;=\; I(\mathbf{r}_{\mathrm{TOA}}, \mathbf{\Omega}) T_{\mathbf{r}_1, \mathbf{\Omega}}(s_{\mathrm{TOA}}), \tag{4.32}$$

with the transmission function

$$T_{\mathbf{r}, \mathbf{\Omega}}(s) \;=\; \exp\left(- \int_0^s \beta_e(s')\, ds' \right). \tag{4.33}$$

along a characteristic line through point $\mathbf{r}$ in direction $\mathbf{\Omega}$. Here, s denotes the pathlength along the line with respect to point $\mathbf{r}$. Hence, s_{TOA} in (4.32) represents the full pathlength from point $\mathbf{r}_1$ till the upper model boundary. Equation (4.32), together with the boundary condition (4.31), allows the calculation of the direct component I_{dir} in a straightforward manner.

The diffuse component can be expressed by

$$\begin{aligned} I_{\mathrm{diff}}(\mathbf{r}_1, \mathbf{\Omega}) \;=\; & I_{\mathrm{diff}}(\mathbf{r}_2, \mathbf{\Omega}) T_{\mathbf{r}_1, \mathbf{\Omega}}(s_2) \\ & + \int_0^{s_2} ds' T_{\mathbf{r}_1, \mathbf{\Omega}}(s') [J_{\mathrm{msc}}(\mathbf{r}(s'), \mathbf{\Omega}) + J_{\mathrm{ssc}}(\mathbf{r}(s'), \mathbf{\Omega})]. \end{aligned} \tag{4.34}$$

Here, s_2 is the pathlength between the points $\mathbf{r}_1$ and point $\mathbf{r}_2$. The multiple and the single scattering source function J_{msc} and J_{ssc}, respectively, in (4.34) are

given by

$$J_{\mathrm{msc}}(\mathbf{r}(s),\mathbf{\Omega}) = \frac{\beta_s(r)}{4\pi}\int\limits_{4\pi} d\tilde{\Omega}P(r,\tilde{\mathbf{\Omega}},\mathbf{\Omega})I_{\mathrm{diff}}(\mathbf{r},\tilde{\mathbf{\Omega}}), \tag{4.35}$$

$$J_{\mathrm{ssc}}(\mathbf{r}(s),\mathbf{\Omega}) = \frac{\beta_s(r)}{4\pi}P(r,\mathbf{\Omega}_0,\mathbf{\Omega})I_{\mathrm{dir}}(\mathbf{r},\mathbf{\Omega}). \tag{4.36}$$

In contrast to the direct component, the diffuse intensity field from (4.34) has to be solved with a numerical radiative transfer model and needs therefore further discussion. Equation (4.34) represents an integral equation for the diffuse intensity field, which can be solved iteratively using the integrations of a representive set of characteristic lines. For the full three-dimensional radiative transfer problem without any symmetry properties, the required number of characteristic lines is large, which hampers any numerical implementation. However, the azimuthal symmetry for the radiative transfer problem at hand simplifies the numerical solution significantly. The intensities at point $\mathbf{r}_1$ and $\mathbf{r}_2$ in (4.34) can be identified with those in the $\Phi = 0$ plane, where the path integrals have to be determined along a three-dimensional path, which starts in the $\Phi = 0$ plane at $\mathbf{r}_1$ in direction $\mathbf{\Omega}$ and ends at the intersection point with a spherical shell of radius $|\mathbf{r}_2|$. This allows one to build up an iteration scheme based on the intensity in the $\Phi = 0$ plane only.

For a numerical implementation we discretize the model atmosphere in N equidistant, optically homogeneous spherical shells with radii r_n $(n = 0,\ldots,N)$. In the global zenith direction we use a discretization in M different axes Ψ_m $(m = 1,\ldots,M)$. This yields a discretization of the atmosphere in elements, where the corners of one cell are denoted by the index pairs (m,n), $(m+1,n)$, $(m+1,n-1)$ and $(m,n-1)$ (see Fig. 4.3).

Furthermore, the directional dependence of the radiation field in each of these grid points is discretized by $K = K_1 \cdot K_2$ characteristic lines, where K_1 is the number of local zenith directions, θ_{k_1} $(k_1 = 1,\ldots,K_1)$, which are equally spaced in $\mu_{k_1} = \cos\theta_{k_1}$ and K_2 represents an even number of equally spaced streams for the local azimuth directions, φ_{k_2} $(k_2 = 1,\ldots,K_2)$. For such a discretization the integration of a function f over the solid angle can be approximated by

$$\int d\Omega f(\mathbf{\Omega}) = \sum_{k=1}^{K} w_k f(\mathbf{\Omega}_k) \tag{4.37}$$

with the weight $w_k = 4\pi/K$.

Due to the solar illumination which is symmetric around the global zenith, it suffices to consider only one hemisphere in the radiative transfer calculations as depicted in Fig. 4.3. However, this introduces the global zenith axis at $\Psi_1 = 0°$ and $\Psi_M = 180°$ as a new model boundary into the radiative transfer problem. In order to reduce the discretized domain, the model boundary in the shadow region can be shifted towards the edge of the shadow zone, as any radiative contribution from outside of the boundary will be almost zero. Therefore, the new model boundary can be constituted at $\Psi_M = \Psi_{\mathrm{shadow}}$. At the model boundaries appropriate boundary conditions have to be imposed. Fig. 4.3 illustrates the individual boundary conditions. Incoming radiation is defined as entering an atmospheric element from outside the discretized domain. Accordingly, outgoing radiation leaves

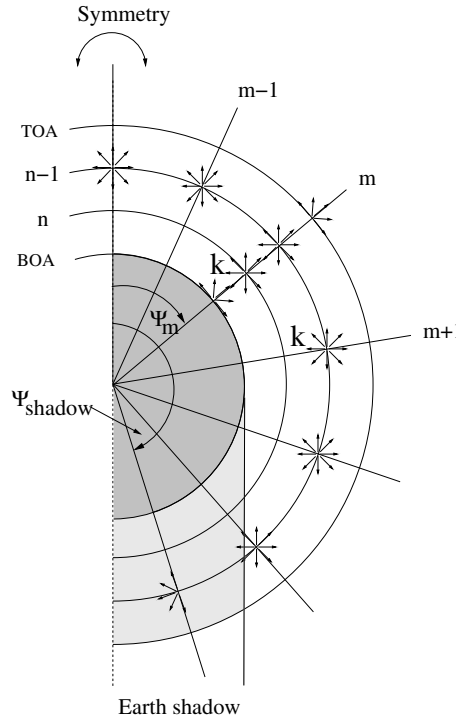


Fig. 4.3. Illustration of the model atmosphere, which is discretized in N spherical shells ($n = 0, \ldots, N$) and $m = 1, \ldots, M$ zenithal axes Ψ_m. One atmospheric element can be characterized by its corner points (m, n), $(m + 1, n)$, $(m + 1, n - 1)$ and $(m, n - 1)$. The radiation field at a point (m, n) can be described by $k = 1, \ldots, K$ streams, which provide a discretization of the full solid angle. Streams with same indices k at different locations point in the same global direction. Also appropriate boundary conditions are indicated in the sketch. There is no incoming diffuse radiation at the top of the atmosphere (TOA). At the lower boundary (BOA) surface reflection may be included. Further, the radiation field is symmetric around the global zenith. In the shadow region for $\Psi > \Psi_{\text{shadow}}$ the boundary condition is approximated by no incoming radiation. After Walter et al. (2006), [59].

the atmospheric element. At TOA and for all axes Ψ_m no incoming diffuse radiation is allowed. At BOA the effects of a reflecting surface can be included, otherwise incoming radiation is set to zero. The zenith boundary condition at $\Psi_1 = 0°$ is characterized by a symmetric radiation field at all levels n; outgoing radiation in direction $\mathbf{\Omega}_k = (\theta_{k_1}, \varphi_{k_2})$ is equal to incoming radiation in the mirrored direction $\mathbf{\Omega}_{k'} = (\theta_{k_1}, \varphi_{k'_2})$ with $\varphi_{k'_2} = \varphi_{k_2} + 180°$. In contrast, the shadow region at Ψ_{shadow} is subject of no incoming radiation.

For a further evaluation of the diffuse intensity field $I_{\text{diff}}(m, n, k)$ at a grid point (m, n) in the direction of a stream k with (4.34) we separate its single scattering (ssc) and multiple scattering (msc) contribution,

$$I_{\text{diff}}(m, n, k) \;=\; I_{\text{ssc}}(m, n, k) + I_{\text{msc}}(m, n, k). \tag{4.38}$$

Here, the single scattering solution is given by

$$I_{\mathrm{ssc}}(m,n,k) \;=\; \int_0^s ds' \frac{\beta_s(s')}{4\pi} P_{\odot k}(s') F_0 T_{mn\odot}(s_0(s')) T_{mnk}(s'), \qquad (4.39)$$

where $P_{\odot k}(s)$ describes the scattering of photons from the solar direction $\odot$ into direction k at a position on the characteristic line, which is given by the pathlength s. T_{mnk} is the discretized form of the transmission function (4.33). The solar pathlength $s_0(s)$ through the atmosphere, and hence the transmission function $T_{mn\odot}$, depend on the location on the characteristic line s. For the evaluation of (4.39) we integrate the solar contribution along the full path of the characteristic line in direction k, i.e. until a model boundary is reached. This corresponds to a so-called long characteristic method [13,20]. In this manner, the single scattering part $I_{\mathrm{ssc}}(m,n,k)$ can be calculated at every grid point in the atmosphere in every direction. The single scattering contribution is needed to initialize the calculation of the multiple scattering part.

For the evaluation of the multiple scattering part we follow a different strategy. Here we relate the diffuse intensity $I_{\mathrm{diff}}(m,n,k)$ at grid point (m,n) in direction k to the intensity at the intersection point $\mathbf{r}_p$ of the corresponding characteristic line with the spherical shell above or below that grid point using (4.34) (see Fig. 4.4). In general, the intersection point $\mathbf{r}_p$ cannot be assigned to one of the regular grid points of the chosen discretization. Therefore, we determine $I_{\mathrm{diff}}(\mathbf{r}_p,k)$ via a linear interpolation between its neighbouring values defined on the discretization grid. In contrast to (4.39) this approach corresponds to a so-called short characteristic method [13,20], and offers the possibility of an efficient calculation of the multiple scattering radiation field. Furthermore, we assume that the vertical discretization of the model atmosphere is fine enough, so that we can approximate the diffuse intensity field in (4.35) by its layer average

$$\overline{I_{\mathrm{diff}}}(m,n,k) \;=\; \frac{1}{2}\left[I_{\mathrm{diff}}(m,n,k) + I_{\mathrm{diff}}(\mathbf{r}_p,k)\right]. \qquad (4.40)$$

Thus, the multiple scattered contribution of the diffuse intensity field at grid point (m,n) in direction k can be approximated by

$$\begin{aligned} I_{\mathrm{msc}}(m,n,k) \;\approx\; & I_{\mathrm{msc}}(\mathbf{r}_p,k) T_{mnk}(s_p) \\ & + \sum_{\tilde{k}=1}^{K} w_{\tilde{k}} \frac{\beta_s^n}{4\pi} P_{\tilde{k},k}^n \overline{I_{\mathrm{diff}}}(m,n,\tilde{k}) \int_0^{s_p} ds' T_{mnk}(s'). \end{aligned} \qquad (4.41)$$

Here, s_p is the pathlength from grid point (m,n) to the intersection point $\mathbf{r}_p$. The multiple scattering part $I_{\mathrm{msc}}(\mathbf{r}_p,k)$ is again determined via a linear interpolation between the neighbouring values. β_s^n denotes the constant scattering coefficient in the spherical shell with radius r_n and $P_{\tilde{k},k}^n$ denotes the corresponding scattering phase function for photon scattering from direction $\tilde{k}$ into direction k. Equations (4.38)–(4.41) in concert with the boundary conditions enable one to compute the multiple-scattering radiance field by an iterative approach. First the radiative transfer model is initialized with the precalculated single scattering radiance field. Subsequently, the two-dimensional Picard iteration method starts with calculating the multiple scattered radiance with the help of (4.41) in all directions k at the first

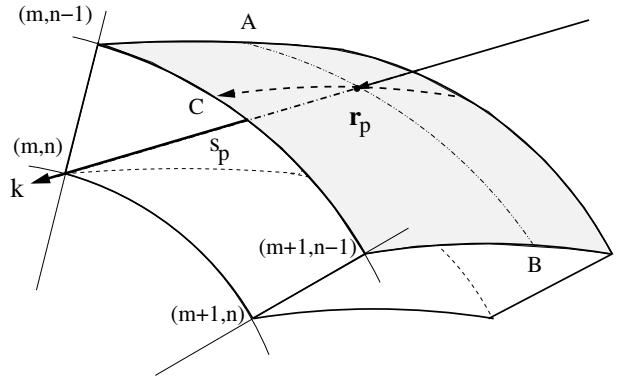


Fig. 4.4. Section of a three-dimensional spherical shell. The characteristic line for downward direction k crosses the spherical shells in the points at (m, n) and $\mathbf{r}_p$. The pathlength s_p between these crossing points is calculated taking the three-dimensional geometry into account. In the short characteristics model for the multiple-scattering computation, any radiative quantity at $\mathbf{r}_p$ in direction k can be calculated by a linear interpolation of the neighbouring values at points A and B in the same direction. Due to the imposed symmetry properties it follows that the radiation field at point A and B is equal to the radiation field at $(m, n-1)$ and $(m+1, n-1)$, which are located in the $\Phi = 0$ plane. In that manner all radiative quantities can be mapped to the $\Phi = 0$ plane, and the linear interpolation for the determination of the radiative quantity at $\mathbf{r}_p$ ultimately takes place at position C between gridpoints $(m, n-1)$ and $(m+1, n-1)$.

axis Ψ_m $(m = 1)$ at the boundary of the uppermost spherical shell $n = 0$ (TOA) and proceeds till the lower boundary at $n = N$.

Then, the corresponding values at all levels for the neighbouring axis Ψ_{m+1} are calculated, taking the updated values of the preceding axis into account. This operation proceeds until the lower boundary of the last axis in the shadow region Ψ_{shadow} is reached. The whole procedure corresponds to one iteration step. The iterations are repeated until the radiation field has converged. For details regarding the Picard iteration approach we refer to [42] and [13].

4.4.1.2 Verification of the forward radiative transfer model

In order to verify the proposed radiative transfer model we compare calculated radiances at the top of the model atmosphere with those calculated by a Monte Carlo reference code. General information on the concept of Monte Carlo calculations and their application in radiative transfer simulations are given by e.g. [29] and [30]. The *McScia* Monte Carlo reference code [50], which has been used for the following comparisons, is based on a backward formulation and takes the full spherical geometry of the radiative transfer problem into account. In a backward Monte Carlo approach, the radiance is calculated by simulating photon trajectories travelling from the detector to the sun. The statistical information related to the individual photon paths allows one to estimate the statistical error of the simulation. *McScia* is the first radiative transfer model, which uses the equivalence theorem [37], to separate scattering from absorption for limb measurements. The *McScia*

Monte Carlo code compares well to other 3-D limb codes for different observational scenarios and is described in detail by [50].

For all simulations we assume a model atmosphere of 100 km geometrical thickness, which has been divided into spherical shells of 1 km vertical extension, thus $N = 100$. The global zenith direction has been discretized in $M = 50$ axes Ψ_m. A total of $K = 48$ streams ($K_1 = 8, K_2 = 6$) were used for the representation of the directional dependence of the radiation field. The computations are performed for observation geometries equal to those used in [27] in order to enhance comparability with other spherical radiative transfer codes. Thus, we assume the satellite instrument to be located directly at the top of the model atmosphere and all angles are given at the tangent point of the satellite's line-of-sight. The US-Standard atmosphere [34] has been used for the simulations. Only Rayleigh scattering and ozone absorption have been taken into account. The calculations were performed for two different wavelengths at 325 nm and 345 nm. The first case represents a scenario with moderate ozone absorption ($\sigma_{\text{abs}} = 1.451 \cdot 10^{-20}$ cm^2, $\sigma_{\text{sca}} = 4.022 \cdot 10^{-26}$ cm^2), whereas the second case corresponds to no absorption but still a relatively large amount of scattering ($\sigma_{\text{sca}} = 3.120 \cdot 10^{-26}$ cm^2).

Fig. 4.5 shows a comparison between the proposed radiative transfer model and a Monte Carlo reference simulation for no surface reflection and an Earth radius of $r_{\text{TOA}} = 6371$ km. The solar zenith angle is 15° and the viewing azimuth angle is 90°. That is, the satellite's line-of-sight forms a 90° angle with the solar illumination. The radiance I^{TOA} is normalized to the incoming solar flux F_0. The graphs on the left-hand side show the reflected radiance as a function of tangent height for two different wavelengths (upper 325 nm, lower 345 nm).

The increase of the reflected radiance with lower tangent heights can be seen clearly. As the lower parts of the atmosphere are optically thicker, the large amount of scattering leads to the observed increase. The graphs on the right-hand side display the relative difference of the Picard iterative method towards the Monte Carlo reference calculations as a function of tangent height. For both wavelengths the differences are smaller than ±1%, and for most tangent heights the deviations are within or just outside the standard deviation of the Monte Carlo computations.

Fig. 4.6 (325 nm) and 4.7 (345 nm) give an overview of the relative difference between the proposed numerical approach and the Monte Carlo model for different viewing and solar geometries. The solar zenith angles are 80° and 90°. In addition to an azimuth angle of 90° also azimuth angles of 20° and 160° are considered. In the first case the satellite's line-of-sight points towards the sun with a 20° offset, whereas in the second case the line-of-sight points away from the sun with the corresponding 160° offset. In general, the agreement of the Picard iterative method with the Monte Carlo computations is good and differences stay below ±3% for all tangent heights. The only exception occurs at a solar zenith angle of 90° for both wavelengths. Here, for the azimuth angles of 20° and 160° relative errors of up to −5% can be observed. For these cases the line of sight crosses atmospheric regions in the vicinity of the shadow region, where a strong gradient in the radiation field is present. The ability of the radiative transfer model to resolve this gradient, strongly depends on the employed discretization. Thus, the observed deviations can be explained by a too rough gridding in this atmospheric region. Further, the larger relative errors can be explained by the circumstance that the absolute values

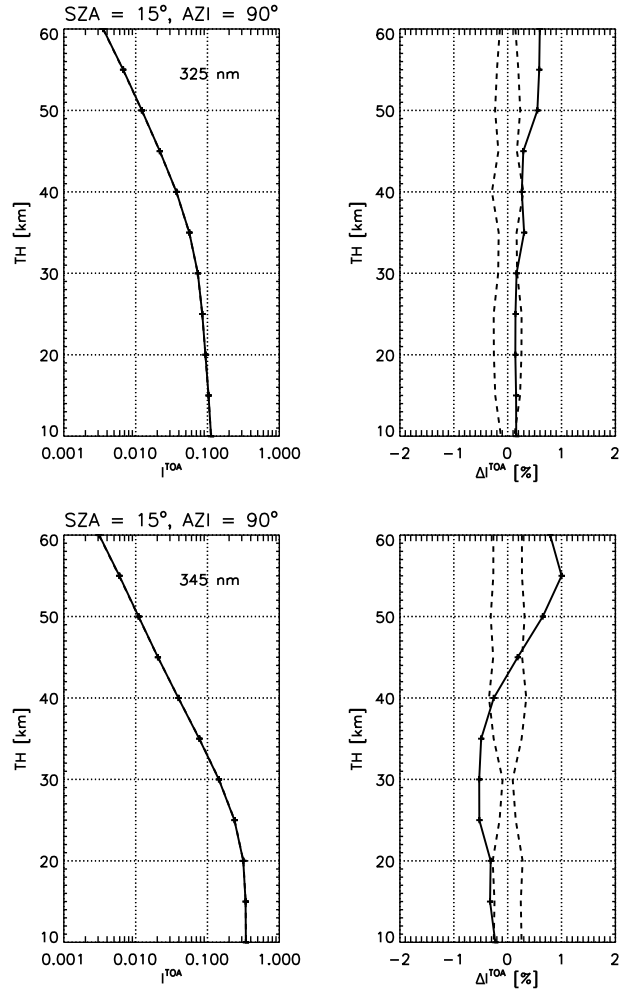


Fig. 4.5. Upper left: Radiance at the top of the model atmosphere I^{TOA} normalized to the incoming solar flux as a function of tangent height (TH). The calculations are performed with the proposed numerical model for a wavelength of 325 nm. The solar zenith angle (SZA) is 15°, the viewing azimuth angle (AZI) is 90°, both are defined at the tangent point of the satellite's line-of-sight. Upper right: Relative difference ΔI^{TOA} of the Picard iterative method with respect to the Monte Carlo reference calculation as a function of tangent height (solid line). The dashed lines represent the standard deviation of the Monte Carlo computations. After Walter et al. (2006), [59]. Lower: same as before, but for a wavelength of 345 nm.

of the backscattered radiance are very small in these cases. Therefore, already small absolute deviations lead to big relative errors.

The presented results are similar to those reported in [27]. With the verification conducted in this section, we conclude that the proposed spherical radiative transfer model is well-suited for calculating the radiation field with sufficient accuracy.

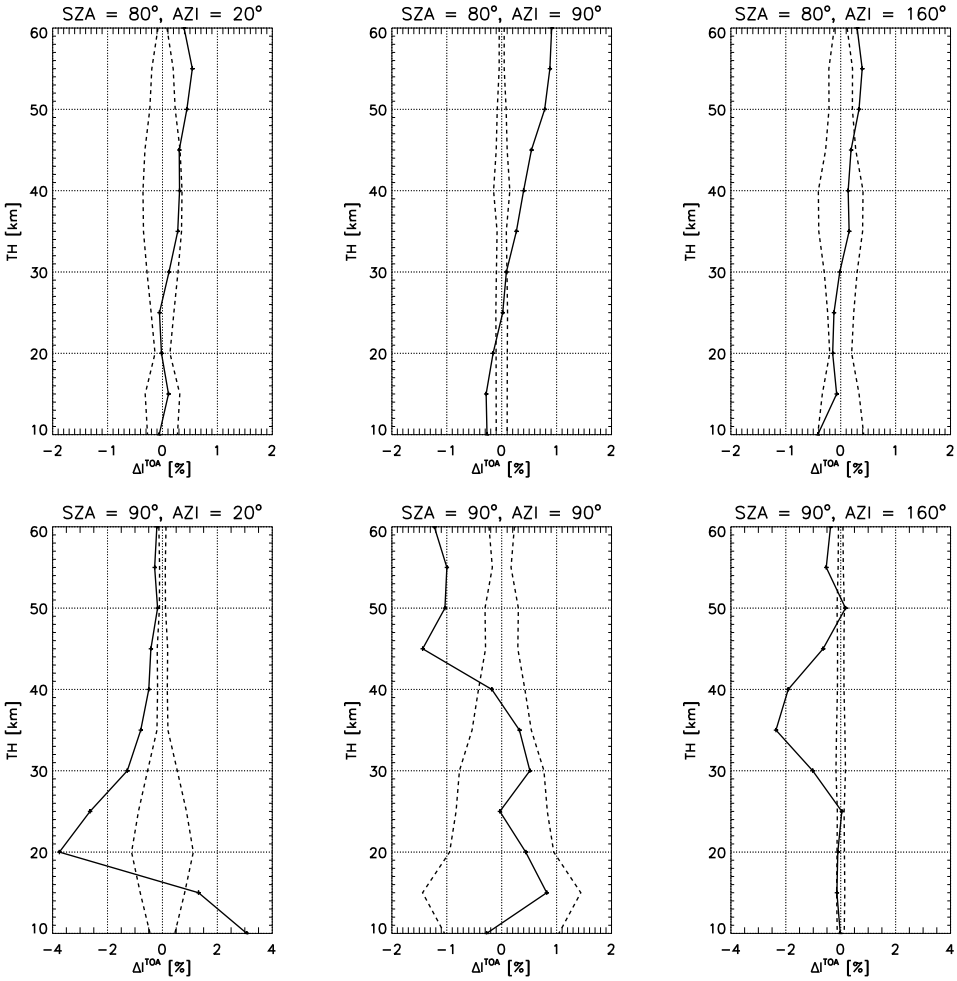


Fig. 4.6. Relative difference (solid line) of the proposed numerical method compared to the Monte Carlo reference calculations for different solar zenith angles (80° and 90°) and viewing azimuth angles (20°, 90° and 160°). The dashed lines represent the standard deviation of the Monte Carlo computations. All calculations are performed for a wavelength of 325 nm. After Walter et al. (2006), [59].

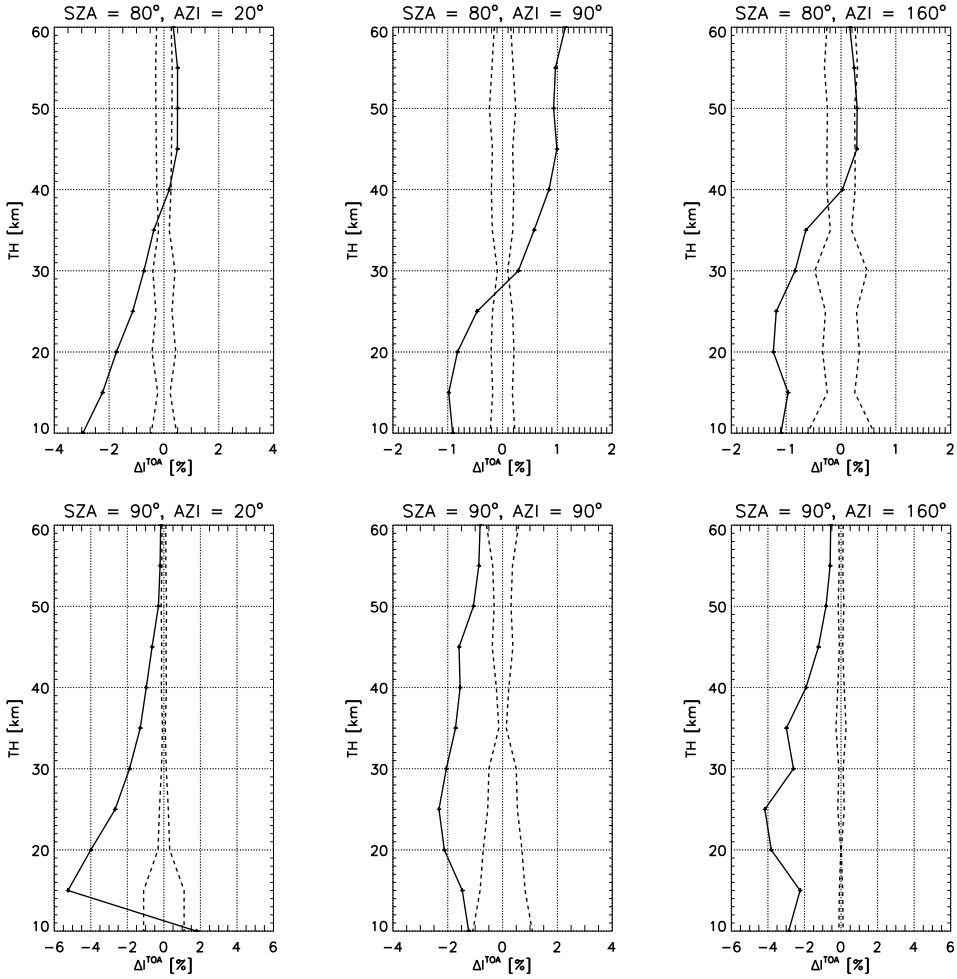


Fig. 4.7. Same as Fig. 4.6, but for a wavelength of 345 nm.

4.4.2 Solution of the pseudo-forward transfer equation

In order to calculate the derivatives of the reflected radiance using (4.25), one has to solve the radiative transfer problem in its pseudo-forward formulation (4.19) in addition. The approach to solve the pseudo-forward problem is similar to that of the forward problem, which was discussed in the previous section, with the only difference in the initialization of the diffuse radiation field. For it we use the equivalent response function R^* from (4.28) as radiation source, taking the needed reversal of directions into account. Analogous to the forward problem we then shift this radiation source into corresponding boundary conditions, which provides the homogeneous equation

$$\mathbf{L}\xi \;=\; 0 \tag{4.42}$$

with the boundary conditions

$$\xi(\mathbf{r}_{\mathrm{TOA}}, \mathbf{\Omega}) \;=\; \frac{1}{\cos\theta_{\mathrm{v}}} \frac{1}{2\pi r_{\mathrm{TOA}}^2 \sin\Psi_{\mathrm{v}}} \delta(\mathbf{\Omega} - \mathbf{\Omega}_{\mathrm{v}}) \frac{1}{d\Psi_{\mathrm{v}}} \tag{4.43}$$

for $-1 \leq \mu < 0$ and $\Psi_{\mathrm{v}} < \Psi < \Psi_{\mathrm{v}} + d\Psi_{\mathrm{v}}$, otherwise zero, and

$$\xi(\mathbf{r}_{\mathrm{BOA}}, \mathbf{\Omega}) \;=\; 0 \tag{4.44}$$

for $0 < \mu \leq 1$. In (4.43) we expressed the Dirac delta function by a Boxcar function of infinitesimally small width $d\Psi_{\mathrm{v}}$, viz.

$$\delta(\Psi - \Psi_{\mathrm{v}}) \;=\; \frac{1}{d\Psi_{\mathrm{v}}} \left[\Theta(\Psi - \Psi_{\mathrm{v}}) - \Theta(\Psi - (\Psi_{\mathrm{v}} + d\Psi_{\mathrm{v}}))\right]. \tag{4.45}$$

Analogous to the forward case, a splitting of the pseudo-forward intensity field ξ in its direct and diffuse components ξ_{dir} and ξ_{diff} leads to two coupled differential equations for the direct beam and the diffuse radiation field. An integration over the pathlength of a characteristic line yields for the direct component

$$\xi_{\mathrm{dir}}(\mathbf{r}_1, \mathbf{\Omega}) \;=\; \xi(\mathbf{r}_{\mathrm{TOA}}, \mathbf{\Omega}) T_{\mathbf{r}_1, \mathbf{\Omega}}(s_{\mathrm{TOA}}) \tag{4.46}$$

with s_{TOA} the full pathlength. This solution can be thought to consist of many satellite lines-of-sight, which are rotated around the global zenith. It therefore constitutes a one-sheeted hyperboloid (see Fig. 4.8).

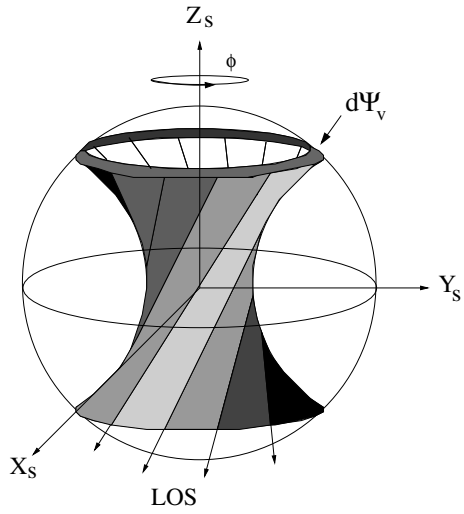


Fig. 4.8. One-sheeted hyperboloid on which the direct component ξ_{dir} of the pseudo-forward intensity field is defined. It can be thought to consist of a rotation Φ of the satellite's line-of-sight (LOS) around the global zenith at Z_s. The actual shape of the hyperboloid depends on the local viewing angles θ_{v} and φ_{v} of the satellite instrument. The thickness of the hyperboloid is given by the infinitesimal small width $d\Psi_{\mathrm{v}}$. After Walter et al. (2006), [59].

The formal solution for the pseudo-forward diffuse radiation field is given by

$$\begin{aligned}\xi_{\text{diff}}(\mathbf{r}_1, \mathbf{\Omega}) &= \xi_{\text{diff}}(\mathbf{r}_2, \mathbf{\Omega}) T_{\mathbf{r}_1, \mathbf{\Omega}}(s_2) \\ &\quad + \int_0^{s_2} ds' T_{\mathbf{r}_1, \mathbf{\Omega}}(s') \left[J^*_{\text{msc}}(\mathbf{r}(s'), \mathbf{\Omega}) + J^*_{\text{ssc}}(\mathbf{r}(s'), \mathbf{\Omega}) \right]. \quad (4.47)\end{aligned}$$

Here, J^*_{msc} and J^*_{ssc} denote the pseudo-forward source function for the multiple and single scattering intensity, respectively. They are defined analogous to (4.35) and (4.36).

Again, the diffuse pseudo-forward intensity may be split in its multiple and single scattering contribution, respectively. For the calculation of the pseudo-forward single scattering contribution $\xi_{\text{ssc}}(m, n, k)$ we determine the intersections of each characteristic line k with the hyperboloid. Depending on position (m, n) and direction k either no, one or two intersections may occur (see Fig. 4.9). At the intersection points the single scattering source function provides a contribution along the infinitesimal small pathlength elements ds_c. Taking only intersections into account that are located within the atmosphere, yields

$$\begin{aligned}\xi_{\text{ssc}}(m, n, k) &= \frac{1}{\cos\theta_{\text{v}}} \frac{1}{2\pi r^2_{\text{TOA}} \sin\Psi_{\text{v}}} \sum_{c=1}^{c_{\max}} \frac{ds_c}{d\Psi_{\text{v}}} \frac{\beta_s(s_c)}{4\pi} P_{vk}(s_c) \\ &\quad T_{mnv}(l(s_c)) T_{mnk}(s_c), \quad (4.48)\end{aligned}$$

where s_c is the pathlength characterizing an intersection point of the line-of-sight with the hyperboloid and $c_{\max} = 0, 1, 2$ indicates the number of intersections. Further, P_{vk} describes the scattering of photons from the satellite's viewing direction $\mathbf{\Omega}_{\text{v}}$ into direction k. In (4.48) v indicates the viewing direction $\mathbf{\Omega}_{\text{v}}$. The pathlength along the satellite's line-of-sight $l(s_c)$ ranges from TOA to the intersection point at s_c. The derivative $ds_c/d\Psi_{\text{v}}$ in (4.48) describes the dependence of the pathlength element ds_c, along which the direct radiation is scattered in direction k, on the width of $d\Psi_{\text{v}}$ in the boundary condition (4.43). We determine the derivative $ds_c/d\Psi_{\text{v}}$ numerically by approximating it to its differential quotient $\Delta s_c/\Delta\Psi_{\text{v}}$. Here, we consider two one-sheeted hyperboloids, separated by $\Delta\Psi_{\text{v}}$ and determine the pathlength Δs_c along an intersecting line-of-sight (see Fig. 4.9). Finally, the single scattering contribution (4.48) is used to initialize the radiative transfer model for the calculation of the multiple scattered radiation field. The calculation of the latter is performed analogously to the calculation of the multiple scattered forward radiation field, as described in subsection 4.4.1.1.

4.4.3 Verification of the adjoint radiation field

Once the pseudo-forward radiance field $\xi(\mathbf{r}, \mathbf{\Omega})$, and therefore the adjoint intensity field $I^\dagger(\mathbf{r}, -\mathbf{\Omega})$, has been calculated, their accuracy can be verified with the help of (4.16), i.e. the alternative evaluation of the radiative effect, which is in our case the radiance at the top of the atmosphere

$$\begin{aligned}I^{\text{TOA}} &= \langle I^\dagger | S_0 \rangle \\ &= 2\pi r^2_{\text{TOA}} \int d\Psi \sin\Psi \cos\Psi F_0 I^\dagger(r_{\text{TOA}}, \Psi, \mathbf{\Omega}_0), \quad (4.49)\end{aligned}$$

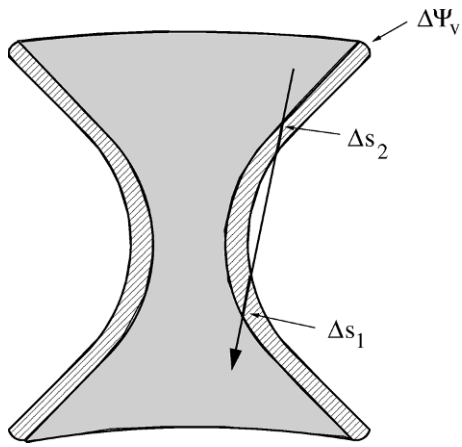


Fig. 4.9. Section of the hyperboloid. Shown are the intersections of an internal characteristic line with the hyperboloid, which are needed for the calculation of the pseudo-forward single scattering contribution. At these points, the illuminated pathlength elements Δs_1 and Δs_2 are determined numerically. They are proportional to the angular distance $\Delta\Psi_{\mathrm{v}}$.

where we made use of the fact, that $\mu_0 = \cos\theta_0 = \cos\Psi$. For the discretized model atmosphere, (4.49) can be approximated by a summation over all axes Ψ_m

$$I^{\mathrm{TOA}} \;\approx\; 2\pi r_{\mathrm{TOA}}^2 \sum_{m=2}^{M} \Delta\Psi_m \sin\overline{\Psi}_m \cos\overline{\Psi}_m F_0 \overline{I}^{\dagger}(m,0,\odot), \tag{4.50}$$

where $\overline{\Psi}_m = (\Psi_{m-1} + \Psi_m)/2$ denotes the averaged global zenith angle and $\Delta\Psi_m = \Psi_m - \Psi_{m-1}$. The averaged adjoint intensity $\overline{I}^{\dagger}(m,0,\odot)$ at the top of the model atmosphere in solar direction is given by

$$\overline{I}^{\dagger}(m,0,\odot) \;=\; \frac{1}{2}(I^{\dagger}(m-1,0,\odot) + I^{\dagger}(m,0,\odot)). \tag{4.51}$$

The calculation of the individual contributions $I^{\dagger}(m,0,\odot)$, $m = 1,\ldots,M$, takes place analogously to Eqs (4.38)–(4.41), i.e. an integration of the diffuse adjoint source function along the corresponding characteristic line in solar direction. Therefore, any error in the calculation of the internal adjoint intensity field will immediately affect the accuracy of the calculated radiance at the top of the atmosphere according to (4.50).

Table 4.1 shows a comparison of the radiance I^{TOA} for different tangent heights calculated with (4.50) and with the forward radiative transfer model for the two wavelengths 325 nm and 345 nm. The discretization of the model atmosphere is the same as in the previous section. The solar zenith angle has been chosen as 60°, due to a significant contribution of multiple scattering to I^{TOA}. The local azimuth angles are 180°, 140° and 100°. For $\varphi_{\mathrm{v}} = 180°$ the satellite's line-of-sight points straight away from the sun. For the two azimuth angles $\varphi_{\mathrm{v}} = 180°$ and 140°, the relative difference ΔI^{TOA} between the two approaches ranges between −2% and +2% for both wavelengths and all tangent heights. Larger deviations appear for a local azimuth angle of 100°. In this case, differences of up to ±8% may occur.

Table 4.1. Verification of the adjoint radiative transfer calculations for the wavelengths 325 nm and 345 nm. The verification is performed at different tangent heights (TH) for a solar zenith angle (SZA) of 60° and three different local azimuth angles $\varphi_v = 180°, 140°$ and 100°. All angles are given at the tangent point. The table shows the the relative difference ΔI^{TOA} between the radiance at the top of the atmosphere computed with the forward radiative transfer model $I^{\mathrm{TOA}}_{\mathrm{FWD}}$ and the adjoint mode $I^{\mathrm{TOA}}_{\mathrm{ADJ}}$.

SZA 60°		325 nm	345 nm
φ_v [°]	TH [km]	ΔI^{TOA} [%]	ΔI^{TOA} [%]
	10	0.2	−1.4
180	25	0.6	−1.0
	50	0.2	1.3
	10	−0.8	−1.6
140	25	1.4	1.1
	50	0.9	1.9
	10	1.3	−1.9
100	25	−7.4	7.9
	50	7.8	6.9

The difference between the adjoint and the forward simulation of I^{TOA} is caused by the particular form of the adjoint radiation source. Here, the single scattering contribution of the adjoint field at position $\mathbf{r}$ in direction $\mathbf{\Omega}$ is only non-zero if the corresponding line-of-sight intersects with the hyperboloid of the direct beam. For the discretized problem this can result in a significant error for the single scattering field in case that the discretization cannot account properly for the spatial extension of the hyperboloid. For example, the number of streams K and by that the number of characteristic lines has a strong influence on the accuracy of the adjoint calculations. If the sterical resolution is small, the hyperboloid volume that is illuminated by ξ_{dir} cannot be resolved properly. This discretization error may lead to an over- or underestimation of the corresponding single scattering contribution. Such effects do not occur for the forward radiative transfer model. Furthermore, Fig. 4.10 shows the diffuse forward and adjoint intensity field, averaged over solid angle, as a function of radius r and zenith angle Ψ. Due to the smooth solar illumination at the top of the atmosphere the forward intensity field is a relatively smooth function, which varies only over one order of magnitude in the relevant spatial domain. In contrast, the pointlike adjoint radiation source leads to an adjoint radiation field which is highly peaked. Depending on the absorption and scattering properties of the atmosphere, the main contribution to the adjoint diffuse intensity due to the adjoint source takes place at or just before the tangent point of the satellite's line-of-sight and then decreases by several orders of magnitude within a relatively small distance around its maximum. To resolve this feature in the diffuse adjoint field a fine spatial discretization in r and Ψ is needed, which is in conflict with an efficient numerical performance of the model. The effect of the described discretization errors becomes larger, the more local and peaked the adjoint field is. Especially, if the local azimuth angle φ_v approaches 90° the hyperboloid starts to deform. This deformation enhances the local character of the adjoint radiation

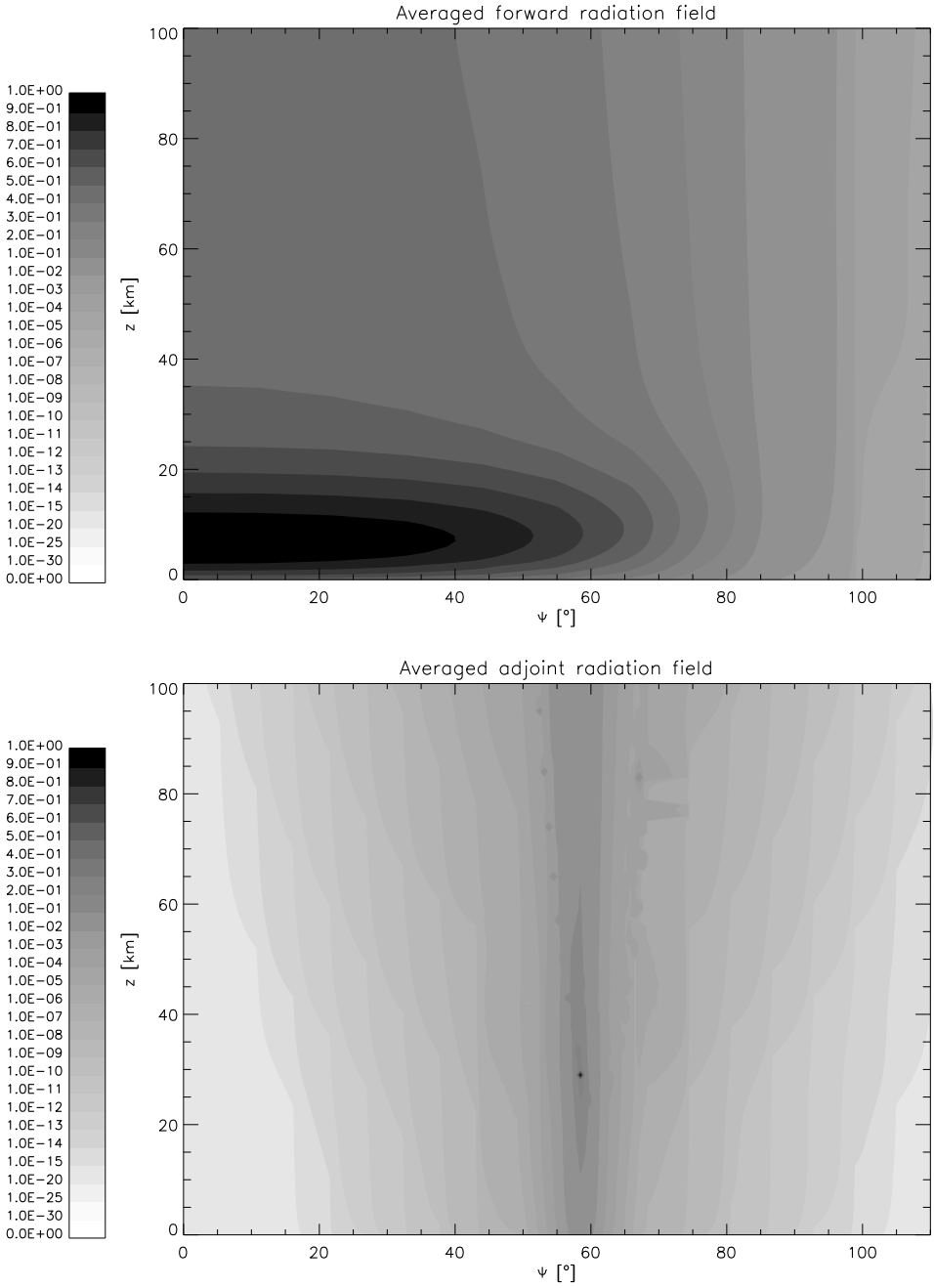


Fig. 4.10. The graphs show the averaged diffuse forward (upper) and adjoint (lower) intensity field $\overline{I}^{\dagger}_{\mathrm{diff}} = \sum_{k=1}^{K} w_k I^{\dagger}_{\mathrm{diff}}(m,n,k)$, respectively. Each intensity field is normalized to its largest value. The simulations are performed for a solar zenith angle of 60°, a local azimuth angle of 140° and a tangent height of 25 km. The wavelength is 325 nm. A discretization of $N = 100$ spherical shells, $M = 50$ zenithal axes Ψ_m and $K = 80$ streams is used. The highly peaked character of the adjoint field is evident.

field, which additionally complicates its proper discretization in all three variables, i.e. $\mathbf{\Omega}$, r and Ψ. This is the main cause for the increased deviations in Table 4.1 for the $\varphi_{\mathrm{v}} = 100°$ case. However, for a very fine discretization these deviations can also be reduced and the result will converge to the solution from the forward calculations.

4.5 Linearization of the spherical radiative transfer model

For a given forward and pseudo-forward radiance field, (4.25) provides an analytical approach for calculating the derivative of the reflected radiance with respect to the trace gas number density x in a model layer n. As the solar source and the equivalent response function, Eqs (4.5) and (4.20), are both defined as surface sources, they do not experience any perturbation for a change in the trace gas number density x within the atmospheric volume and therefore the two last terms of (4.25) vanish.

Therefore, substitution of (4.26) in (4.25), whilst taking (4.18) into account, yields the derivatives for a satellite observation of the radiance at the top of the atmosphere

$$\frac{\partial I^{\mathrm{TOA}}}{\partial x}(x_0) \;=\; -\int\limits_{\Delta V} dV \sigma_{\mathrm{abs}}(\mathbf{r}) \int\limits_{4\pi} d\Omega\; \xi_0(\mathbf{r}, -\mathbf{\Omega}) I_0(\mathbf{r}, \mathbf{\Omega}). \tag{4.52}$$

Due to the symmetry in the global azimuth Φ of the forward and pseudo-forward problem, (4.52) reduces to

$$\begin{aligned}\frac{\partial I^{\mathrm{TOA}}}{\partial x}(x_0) \;=\; & 2\pi \frac{\sigma^n_{\mathrm{abs}}}{2} \int\limits_{r_{n-1}}^{r_n} dr\, r^2 \int\limits_0^{\pi} d\Psi\, \sin\Psi \int_{4\pi} d\Omega \\ & \xi_0(r, \Psi, -\mathbf{\Omega}) I_0(r, \Psi, \mathbf{\Omega}).\end{aligned} \tag{4.53}$$

Here, σ^n_{abs} is the absorption cross-section of the considered trace gas species in a model layer n. The radiance field $I_0(r, \Psi, \mathbf{\Omega})$ and the pseudo-forward field $\xi_0(r, \Psi, -\mathbf{\Omega})$ in (4.53) are the solutions of the forward and the pseudo-forward radiative transfer problem for the unperturbed atmospheric parameter x_0. The perturbation integrals in (4.53) can be evaluated numerically with low computational costs. For this purpose the integral expressions are approximated by corresponding summations of the discretized contributions of the forward and pseudo-forward field, viz.

$$\begin{aligned}\frac{\partial I^{\mathrm{TOA}}}{\partial x}(x_0) \;\approx\; & -2\pi \frac{\sigma^n_{\mathrm{abs}}}{2} \Delta r_n \overline{r}^2_n \left[\sum_{m=2}^{M} \Delta\Psi_m \sin\overline{\Psi}_m \right. \\ & \left. \left(\sum_{k=1}^{K} w_k \overline{\xi}_0(m, n, -k) \overline{I}_0(m, n, k) \right) \right].\end{aligned} \tag{4.54}$$

Here, $\overline{r}_n = (r_{n-1} + r_n)/2$ denotes the averaging over the model layer n and $\overline{\Psi}_m = (\Psi_m + \Psi_{m-1})/2$ the averaging over the zenithal axes m. Further, $\Delta r_n = r_{n-1} - r_n$

and $\Delta\Psi_m = \Psi_m - \Psi_{m-1}$. The quantities $\overline{\xi}_0$ and $\overline{I}_0$ in (4.54) are averaged over the model layer n and over the axis m, respectively. The direction $-k$ is the direct opposite of direction k, thus reversed by 180°.

In order to verify the proposed approach, we choose ozone as the relevant trace gas species to be investigated. For reference purposes we use the finite difference approach

$$\frac{I^{\mathrm{TOA}}(x_0 + \Delta x) - I^{\mathrm{TOA}}(x_0)}{\Delta x} \longrightarrow \left.\frac{\partial I^{\mathrm{TOA}}}{\partial x}\right|_{x_0} \qquad \text{for} \qquad \Delta x \to 0, \tag{4.55}$$

which converges to the derivative for small perturbations in the ozone number density Δx. The perturbed and the unperturbed radiance at the top of the model atmosphere, $I^{\mathrm{TOA}}(x_0 + \Delta x)$ and $I^{\mathrm{TOA}}(x_0)$, are calculated with the forward mode of the spherical radiative transfer model described in subsection 4.4.1.1. The magnitude of the perturbation Δx was chosen in such a manner that a stable behaviour in the numerical calculation of the differential quotient in (4.55) could be achieved. This has been controlled by different calculations with varying values for Δx. Fig. 4.11 shows the derivative $\partial I^{\mathrm{TOA}}/\partial x$ calculated with the spherical perturbation theory approach as well as its relative deviation to the finite difference method for a wavelength of 325 nm. The sensitivity of I^{TOA} with respect to changes

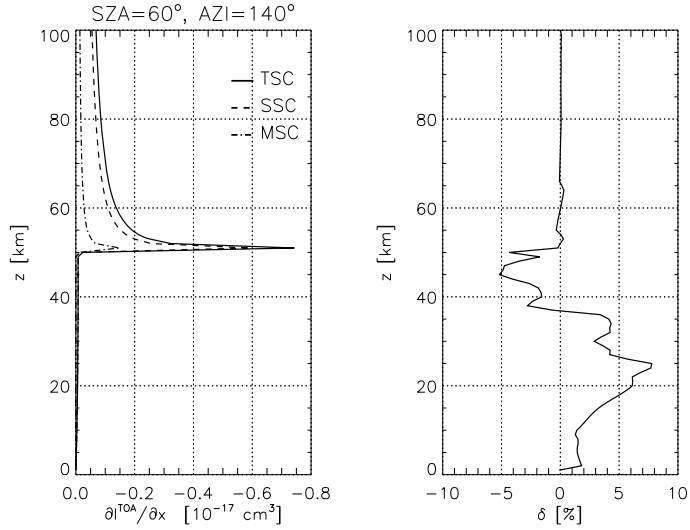


Fig. 4.11. Left panel: Derivative of the reflected radiance I^{TOA} normalized to the incoming solar flux with respect to the ozone density at different altitude levels at 325 nm. The derivative is calculated with the spherical perturbation theory approach from (4.54). The solid line denotes the total scattering contribution (TSC). The dashed line represents the single scattering part (SSC), whereas the dash-dotted line only represents the contribution which results from multiple scattering (MSC). The calculation of the derivative is performed for a tangent height of 50 km, a solar zenith angle of 60° and a local azimuth angle of 140°. The computations are for clear sky conditions. Right panel: Relative difference δ between computed derivatives using the spherical perturbation theory and the finite difference scheme (4.55). After Walter et al. (2006), [59].

in the ozone number density shows an increasing behaviour for decreasing altitudes. Slightly above the tangent height of 50 km a distinctive peak in the sensitivity can be observed. The peak is caused by the very high extinction of radiation along the line-of-sight in the tangent layer. This peak appears in both, the single scattering and the multiple scattering contribution to the sensitivity. In general, the multiple scattering fraction constitutes approximately one-third of the total sensitivity above the tangent height, which demonstrates the necessity to explicitly take this contribution into account in the linearization of the forward model. Below the tangent height the sensitivity is determined by multiple scattering only and drops to very small values for this case.

The relative difference between the perturbation theory approach and the finite difference method is displayed on the right-hand side of Fig. 4.11. Both approaches agree very well in the altitude range from 50 km to 100 km. At the tangent height, just below the maximum sensitivity, the spherical perturbation theory underestimates the reference value by approximately −4%. This is due to the layer averaging property in the perturbation integrals in (4.54), with which it is not possible to take the strong gradient around the tangent height into account. Larger deviations occur below the tangent height of 50 km, where the linearization is based on the multiple-scattering contribution only. In this altitude range, the accompanying radiation fields become very small. As a consequence errors in the calculation of these radiation fields immediately affect the accuracy of the calculated derivatives for the finite difference scheme as well as for the spherical perturbation theory approach. For the latter, discretization errors especially in the calculation of the diffuse adjoint intensity field have a large impact. This can be seen clearly, if one uses the direct-diffuse splitting of the associated radiation fields.

Then, in the approximated perturbation integral, (4.54), products of the diffuse and direct components have to be evaluated, viz.

$$\begin{aligned} \xi_0 I_0 &= (\xi_{\mathrm{diff},0} + \xi_{\mathrm{dir},0})(I_{\mathrm{diff},0} + I_{\mathrm{dir},0}) \\ &= \xi_{\mathrm{diff},0} I_{\mathrm{diff},0} + \xi_{\mathrm{diff},0} I_{\mathrm{dir},0} + \xi_{\mathrm{dir},0} I_{\mathrm{diff},0} + \xi_{\mathrm{dir},0} I_{\mathrm{dir},0}. \end{aligned} \tag{4.56}$$

The last term on the right-hand side generally vanishes, unless the satellite instrument looks directly into the sun. As the component $\xi_{\mathrm{dir},0}$ is always zero below the tangent height, the third term only contributes above the tangent height. It denotes the direct observation of the diffuse radiation field by the satellite and has a relatively large contribution to the perturbation integral. Thus, below the tangent height only the first two terms of (4.56) contribute to the derivative. Both include the diffuse pseudo-forward intensity field $\xi_{\mathrm{diff},0}$, which is affected by discretization errors as discussed in subsection 4.4.3. Therefore errors in the calculation of the diffuse pseudo-forward field will dominate the total error in the calculated derivatives below the tangent height. This leads to errors of up to 8% in the derivative. However, these errors are related to very small sensitivities and are therefore of minor importance in a trace gas profile retrieval.

Figs 4.12 and 4.13 show the same comparison, but for a tangent height of 25 km and 10 km, respectively. In Fig. 4.12 the distinctive peak of the derivatives at the tangent height is again the dominant feature. But also a smoothed sensitivity between approximately 25 km and 40 km can be seen. It results from the relatively large extinction of radiation at the considered wavelength in this atmospheric al-

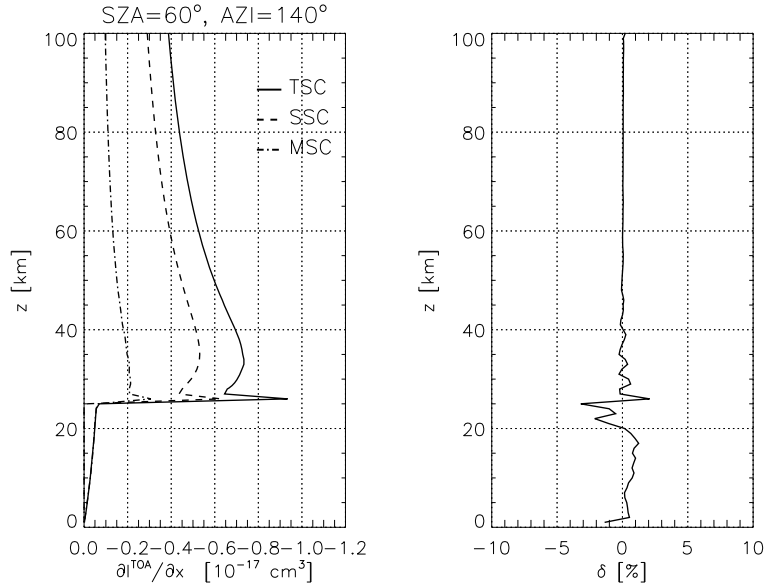


Fig. 4.12. Same as Fig. 4.11, but for a tangent height of $TH = 25\,\mathrm{km}$. After Walter et al. (2006), [59].

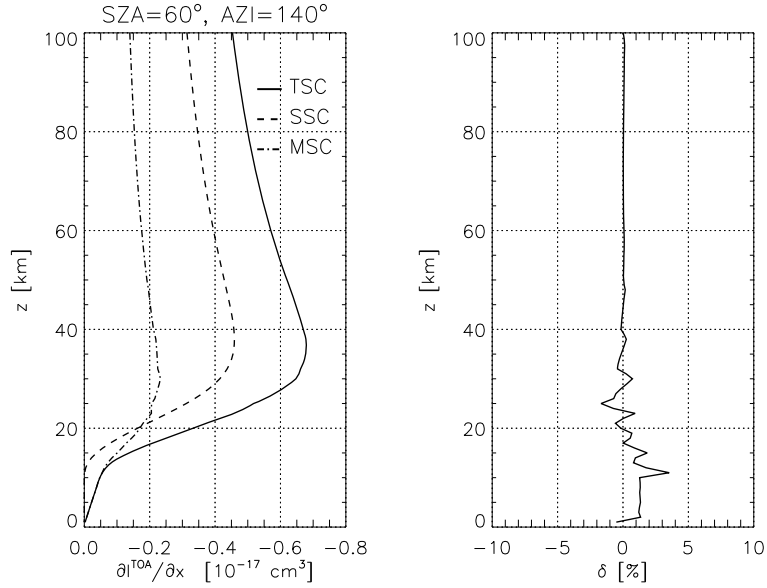


Fig. 4.13. Same as Fig. 4.11, but for a tangent height of $TH = 10\,\mathrm{km}$. After Walter et al. (2006), [59].

titude range. In addition, the multiple scattering contribution to the sensitivity increases as the satellite's line-of-sight senses lower parts of the atmosphere. The behaviour of the relative difference resembles the one from Fig. 4.11. There is a very good agreement above the tangent height and increasing deviations below the tangent layer. Here, relative errors of up to ±3% can be observed. In contrast to Fig. 4.11 and Fig. 4.12, the distinctive peak in the sensitivity disappears for a tangent height of 10 km (Fig. 4.13). Extinction due to absorption and scattering is for this wavelength and altitude range so large, that the smoothing effect dominates the sensitivity. The large pathlength of the line-of-sight in the tangent layer is of no importance anymore, as most radiation already gets attenuated in the layers above. Therefore, if this derivative were to be used in a limb remote sensing application, the vertical resolution of the retrieved trace gas profiles, which depends on the sharply peaked sensitivity, would be rather bad. In such a case it is advisable to shift to another wavelength where less extinction takes place, in order to retain the sharp resolution of a limb weighting function again. Additionally, for altitudes below 20 km multiple scattering effects are more important for the sensitivity than single scattering processes, as can be seen in Fig. 4.13. Here, the relative error remains rather small for all altitudes, as the influence of the discretization error, which originates from the single scattering initialization of the multiple scattering calculation, is reduced by a large number of multiple scattering events. The error reaches a maximum of 4% at the tangent height.

In order to investigate the influence of the spatial and directional discretization on these linearization errors, we calculate derivatives for different model setups. Fig. 4.14 displays the relative difference of the perturbation theory approach with respect to the finite difference method using different discretizations in the number of streams, the number of zenithal axes and the number of spherical shells. In each case we vary the discretization of one variable, whereas we keep the discretization of the other two variables constant. The calculations are performed for the solar and viewing geometry of Fig. 4.12, thus for a solar zenith angle of 60°, an azimuth angle of 140° and a tangent height of 25 km. For the first comparison (left-hand side of Fig. 4.14) we kept the number of spherical shells and the number of zenithal axes constant at $N = 100$ and $M = 50$, respectively, whereas we varied the number of streams, viz. $K = 80, 48, 24$. As expected, with a decreasing number of streams the linearization error becomes bigger. Remarkable is, that even for a very low number of streams the linearization error above the tangent height stays relatively small. For a number of 24 streams only a small offset of approximately 1% can be observed. This again confirms that in this altitude region the main contribution to the derivative is due to the direct observation of the diffuse radiation field, as discussed in connection with (4.56). The direct component of the adjoint field only weakly depends on the used discretization. It is likely, that the observed offset can be attributed to errors in the multiple scattered adjoint field, which is transmitted from lower atmospheric regions. In contrast, larger errors in the calculation of the derivative appear beneath the tangent height. Here, the maximum deviations reach up to 30%. With a decreasing number of streams, the calculation of the associated diffuse adjoint field is dominated by discretization errors. The reason is, as already mentioned before, that the local adjoint radiation source can no longer be resolved properly.

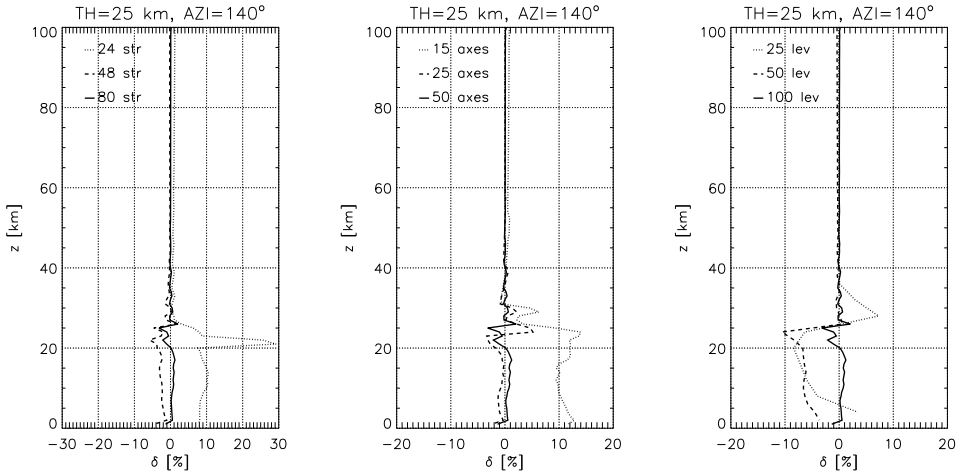


Fig. 4.14. Discretization errors in the calculated derivatives for a solar and viewing geometry as in Fig. 4.12. The graphs show the relative difference of the spherical perturbation theory approach to the finite difference approach. Left: calculations are performed for a different number of streams, viz. 24, 48 and 80 streams. The number of axes for the discretization of the global zenith was kept constant at 50 axes. For the vertical discretization of the model atmosphere, 100 spherical shells were employed. Middle: development of the linearization error for a different number of zenithal axes. The calculations were performed using 15, 25 and 50 axes. Here, a constant number of 80 streams and 100 spherical shells were used. Right: error for a different number of spherical shells, viz. 25, 50 and 100 levels. In these cases 80 streams and 50 zenithal axes were used for the calculations. After Walter et al. (2006), [59].

The graph in the middle of Fig. 4.14 shows the linearization error for a varying number of zenithal axes $M = 50$, 25 and 15. The number of of spherical shells and the number of streams are kept constant at $N = 100$ and $K = 80$, respectively. Here also for a decreasing number of axes, the linearization error becomes bigger. For 25 axes only a small increase in the relative difference can be seen. However, in the case of only $M = 15$ zenithal axes clear deviations appear. Above the tangent height again an offset of approximately 1% can be seen, whereas below the tangent height large deviations occur. Finally, the graph on the right-hand side of Fig. 4.14 shows the evolution of the relative error for different vertical discretizations. Here, we used $N = 100$, 50 and 25 spherical shells, while we kept the number of zenithal axes at $M = 50$ and the number of streams at $K = 80$. With decreasing vertical resolution, inaccuracies in the calculation of the derivatives mainly at and below the tangent height occur. For too coarse a vertical grid, the assumption of a layer averaged diffuse adjoint field according to (4.40) fails, which leads to errors especially in those regions where the vertical gradient of the diffuse radiaton field becomes important, e.g. at the tangent height.

Hence, it is necessary to choose a fine discretization for each of the three investigated variables, i.e. for the directional dependence, the zenith variable and the vertical dimension. With a corresponding finer discretization it is possible to resolve smaller structures in the adjoint field, leading to a more accurate calculation

of the derivatives. Altogether, the investigation of the linearization error showed, that the proposed approach for the calculation of the needed derivatives, although affected by a discretizational bias, is consistent and converges to the exact solution. Discretizational errors appear mainly below the tangent height, where the absolute value of the derivatives is in general very small. Therefore these errors are not expected to negatively influence a trace gas profile retrieval. Thus, despite the observed deviations, it is demonstrated that the spherical perturbation theory approach is suitable for a sufficiently accurate calculation of the derivatives $\partial I^{\mathrm{TOA}}/\partial x$.

Fig. 4.15 shows an example of the calculated derivatives of a limb viewing scenario for different tangent heights at a wavelength of 325 nm. Here, the derivatives of the radiance at the top of the atmosphere with respect to the ozone number density have been determined with (4.54). The narrow peaks of the derivatives at the corresponding tangent heights are the reason for the high vertical resolution that can be achieved in a trace gas profile retrieval from limb measurements. Especially below each individual tangent height only minor contributions to the sensitivity can be seen. In Fig. 4.15 the maximum sensitivity for the considered wavelength appears around an altitude of approximately 35 km. For shorter wavelengths with stronger ozone absorption the corresponding maximum will be shifted

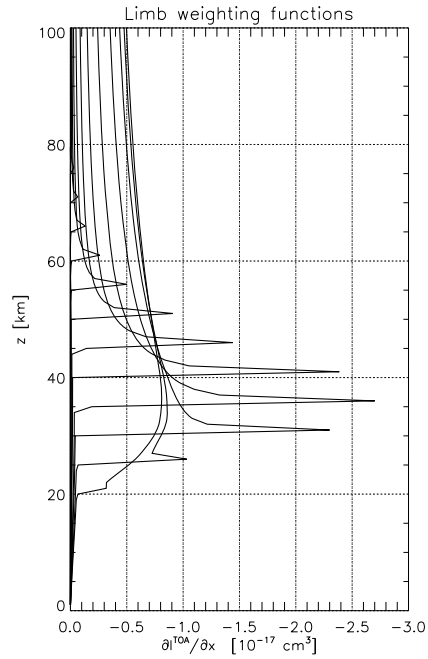


Fig. 4.15. Calculated derivatives or weighting functions of a limb measurement simulation with respect to the ozone number density as a function of height. Here, calculations are performed at a wavelength of 325 nm for tangent heights ranging from 20 km up to 80 km in steps of 5 km. The narrow peaks in the corresponding derivatives enable one to achieve a good vertical resolution in the retrieval of stratospheric ozone profiles.

towards higher altitudes, whereas for longer wavelengths with weaker absorption the maximum appears lower in the atmosphere. Thus, with the vertical sampling and using the wavelength dependence of the corresponding sensitivities it is possible to gain detailed information on the vertical composition of the atmosphere.

4.6 Conclusions

In this study we show that the forward-adjoint perturbation theory approach for radiative transfer in a spherical atmosphere provides a new manner of calculating the derivatives of the backscattered radiance, which are needed in the retrieval of trace gas profiles using satellite limb observations. The proposed perturbation theory approach can be used to linearize any spherical radiative transfer code that calculates the internal forward and adjoint radiation field. In general, the solution of the adjoint problem requires three-dimensional radiative transfer calculations. By the use of an equivalent response function we presented how symmetries of the forward problem can be used to ease the solution of the adjoint problem substantially. Here, we made use of the fact that the response function is not uniquely determined if the forward radiation field is subject to symmetry properties. For example, the assumption of homogeneous spherical shells and a solar illumination, that is symmetric around the global zenith, leads to a forward radiation field that is independent of the global azimuth angle Φ. Therefore the equivalent response function R^* which comprises an integration over the global azimuth angle can be used to describe the same radiative effect E. This equivalent response function R^* represents, in comparison to the unchanged response function R, an adjoint source which illuminates the top of the model atmosphere on a circle centered around the global zenith. Such an illumination introduces symmetries into the adjoint formulation analogous to those of the forward formulation, i.e. the resulting diffuse adjoint field will be independent of the global azimuth angle Φ. Subsequently, this modified adjoint transfer problem can be easily transformed into a pseudo-forward problem, which allows one to use the same computational algorithm as for the calculation of the forward radiation field. In the framework of the forward-adjoint perturbation theory also the impact of a perturbation of internal sources and of the response function onto the derivatives can be taken into account easily. Thus, the forward-adjoint perturbation theory provides a linearization of atmospheric radiative transfer in its most general form.

In order to demonstrate the capability of this method we developed a linearized radiative transfer model for a spherical shell atmosphere. The corresponding radiative transfer equation for the forward mode has been solved by a Picard iteration method. By the use of the equivalent response function it is possible to determine also the adjoint field in a corresponding manner. Here, we describe the limb observation not by a point-like response function directed along the line-of-sight, but by an adjoint source, which is shaped like a one-sheeted hyperboloid. As a result, the adjoint problem reduces to a two-dimensional instead of a three-dimensional radiative transfer formulation in space.

The forward mode of the model is verified with Monte Carlo reference simulations for different solar and viewing geometries at two wavelengths in the ultraviolet

spectral range at 325 nm and 345 nm. The calculated radiances show in general an agreement of approximately ±3%. Only for very large solar zenith angles deviations of up to −5% are observed. The accuracy of the adjoint mode is verified by comparing the limb radiance simulations at the top of the atmosphere, calculated with the adjoint model, with those from the forward mode. Differences are in general below ±2%; however, they can increase up to ±8% for particular viewing geometries. The latter errors are related to an adjoint radiation field with a very local behaviour and therefore discretizational errors in the adjoint model become more important. Hence, the influence of these discretizational errors on the calculation of the adjoint field depends on the observation geometry. The more local the adjoint field is defined, for example if the local azimuth angle approaches 90°, an increasingly fine discretization of the adjoint problem is needed in order to retain a required accuracy.

As surface sources were employed in the description of radiative transfer only the perturbation of the transport operator had to be taken into account in the evaluation of the perturbation integrals. The accuracy of the derivatives of the reflected radiance with respect to the vertical distribution of ozone was investigated for different limb viewing scenarios. The derivatives calculated with the perturbation theory approach were verified with numerical calculations using a finite difference scheme. The calculated derivatives show an excellent agreement above the tangent height of the limb observations. At the tangent height and below, the differences increase. However, it is shown that these errors are mainly due to a deficient directional and spatial discretization of the adjoint intensity field. The errors can be reduced by a finer descretization and therefore do not represent any fundamental restriction of the forward-adjoint perturbation theory approach. For a trace gas profile retrieval the observed differences are of minor relevance, as the sensitivity in the corresponding altitude range is generally at least one order of magnitude smaller than above the tangent height. Therefore the absolute contribution in an actual retrieval is very small.

Appendix A: Transformation of a volume source into a surface source

The ordinary differential equation (ODE) for the intensity, (4.2), is given in volume source notation, viz.

$$\mathbf{L}I \;=\; S_0, \tag{4.57}$$

that is, the solar source S_0 represents a radiation source, which is virtually situated within the atmosphere, just beneath the upper boundary:

$$S_0(\mathbf{r}, \mathbf{\Omega}) \;=\; \mu_0 F_0 \delta(r - r_{\mathrm{TOA}}) \delta(\mathbf{\Omega} - \mathbf{\Omega}_0) \Theta(\mu_0)\,. \tag{4.58}$$

The boundary conditions for the intensity are given by

$$\begin{aligned} I(\mathbf{r}_{\mathrm{TOA}}, \mathbf{\Omega}) \;&=\; 0 \quad \text{for} \quad -1 \leq \mathbf{n} \cdot \mathbf{\Omega} < 0 \\ I(\mathbf{r}_{\mathrm{BOA}}, \mathbf{\Omega}) \;&=\; 0 \quad \text{for} \quad -1 \leq \mathbf{n} \cdot \mathbf{\Omega} < 0, \end{aligned} \tag{4.59}$$

where $\mathbf{n}$ is a unit vector in the direction of the outward normal at position $\mathbf{r}$ on the boundary. Any radiation with $\mathbf{n}\cdot\mathbf{\Omega} < 0$ crosses the boundary in an inward direction, and vice versa. Equation (4.57) is a nonhomogeneous ODE and cannot be solved in a straightforward manner. Therefore we will use the principle of general reciprocity to shift the volume source into a surface source, which subsequently appears in the boundary conditions. This in turn transforms the nonhomogeneous ODE into a homogeneous one, which can be solved easily.

Basically the principle of general reciprocity states, that photon interactions take place in the same sample volume. As this interaction volume does not change when using a different notation, the principle of general reciprocity can be used to relate radiation fields and their sources (volume and surface sources) at different locations in the atmosphere. It is given by [4, 9]

$$\begin{aligned}&\int_A dA \int_{\mathbf{n}\cdot\mathbf{\Omega}<0} d\Omega|\mathbf{n}\cdot\mathbf{\Omega}|\left[I_2(\mathbf{r}_{\mathrm{TOA}},\mathbf{\Omega})I_1(\mathbf{r},-\mathbf{\Omega}) - I_1(\mathbf{r}_{\mathrm{TOA}},\mathbf{\Omega})I_2(\mathbf{r},-\mathbf{\Omega})\right]\\ &= \int_V d\mathbf{r}\int d\Omega\left[I_2(\mathbf{r},-\mathbf{\Omega})S_1(\mathbf{r},\mathbf{\Omega}) - I_1(\mathbf{r},\mathbf{\Omega})S_2(\mathbf{r},-\mathbf{\Omega})\right].\end{aligned} \tag{4.60}$$

Here, I_1 and I_2 are radiation fields, $I_1(\mathbf{r}_{\mathrm{TOA}},\mathbf{\Omega})$ and $I_2(\mathbf{r}_{\mathrm{TOA}},\mathbf{\Omega})$ represent the illumination of the upper boundary and therefore surface sources, whereas S_1 and S_2 are volume sources. $\int_A dA$ denotes the integration over the bounding surface and $\int_V d\mathbf{r}$ represents the integration over the whole volume. From (4.60) several reciprocity principles may be derived, e.g. the well-known optical or Helmholtz reciprocity, which is of great importance in the context of Green's functions [9].

If we set $I_1(\mathbf{r}_{\mathrm{TOA}},\mathbf{\Omega}) = 0$, $S_2 = 0$ and if we identify the remaining I_1, I_2 with the radiation field I, respectively, as well as S_1 with S_0, (4.60) can be used to derive a relation, which provides the transformation of a volume source into a surface source. Thus, we get

$$\begin{aligned}&\int_A dA \int_{\mathbf{n}\cdot\mathbf{\Omega}<0} d\Omega|\mathbf{n}\cdot\mathbf{\Omega}|I(\mathbf{r}_{\mathrm{TOA}},\mathbf{\Omega})I(\mathbf{r},-\mathbf{\Omega})\\ &= \int_V d\mathbf{r}\int d\Omega I(\mathbf{r},-\mathbf{\Omega})S_0(\mathbf{r},\mathbf{\Omega}).\end{aligned} \tag{4.61}$$

It is easy to see, that the two integrations in (4.61) are only equal if

$$I(\mathbf{r}_{\mathrm{TOA}},\mathbf{\Omega}) = \frac{1}{|\mathbf{n}\cdot\mathbf{\Omega}|}S_0(\mathbf{r},\mathbf{\Omega}) = \frac{1}{\cos\theta_0}S_0(\mathbf{r},\mathbf{\Omega}). \tag{4.62}$$

Therefore one is allowed to write

$$\mathbf{L}I \;=\; 0, \tag{4.63}$$

with the new boundary conditions

$$\begin{aligned}I(\mathbf{r}_{\mathrm{TOA}},\mathbf{\Omega}) \;&=\; \frac{1}{\cos\theta_0}S_0(\mathbf{r},\mathbf{\Omega}) \quad \text{for} \quad -1 \le \mathbf{n}\cdot\mathbf{\Omega} < 0\\ I(\mathbf{r}_{\mathrm{BOA}},\mathbf{\Omega}) \;&=\; 0 \quad \text{for} \quad -1 \le \mathbf{n}\cdot\mathbf{\Omega} < 0.\end{aligned} \tag{4.64}$$

Equations (4.63) and (4.64) describe a scenario, in which the boundary is illuminated from outside the atmosphere. The transformation of the volume source into a surface source led to an additional $1/\cos\theta_0$ factor in the boundary conditions. Equations (4.63) and (4.64) form a homogeneous ODE, which can be solved in a straightforward manner by standard mathematical methods.

Acknowledgments

We would like to thank Adrian Doicu from the German Aerospace Center (DLR) for many discussions and helpful suggestions on radiative transfer in a spherical atmosphere. Further we would like to thank Francesco Spada for corresponding Monte Carlo simulations. This research was supported by the Netherlands Institute for Space Research SRON under project number 6430-SCIARALI (GO-2).

References

1. C. N. Adams and G. W. Kattawar. Radiative transfer in spherical shell atmospheres, i. Rayleigh scattering. *Icarus*, **35**:139–151, 1978.
2. D. E. Anderson. The troposphere-stratosphere radiation field at twilight: a spherical model. *Planet. Space Sci.*, **31**:1517–1523, 1983.
3. M. Balluch. A new numerical model to compute photolysis rates and solar heating with anisotropic scattering in spherical geometry. *Ann. Geophysicae*, **14**:80–97, 1996.
4. G. I. Bell and S. Glasstone. *Nuclear Reactor Theory*. Van Nostrand Reinholt, New York, 1970.
5. H. Bovensmann, J. P. Burrows, M. Buchwitz, J. Frerick, S. Noel, V. V. Rozanov, K. V. Chance, and A. P. H. Goede. Sciamachy: Mission objectives and measurement modes. *J. Atmos. Sci.*, **56**:127–150, 1999.
6. M. Box, S. Gerstl, and C. Simmer. Application of the adjoint formulation to the calculation of atmospheric radiative effects. *Beitr. Phys. Atmos.*, **61**:303–311, 1988.
7. M. Box, S. Gerstl, and C. Simmer. Computation of atmospheric radiative effects via perturbation theory. *Beitr. Phys. Atmos.*, **62**:193–199, 1989.
8. M. A. Box, M. Keevers, and B. H. J. McKellar. On the perturbation series for radiative effects. *J. Quant. Spectros. Radiat. Transfer*, **39**:219–223, 1988.
9. K. M. Case. Transfer problems and the reciprocity principle. *Rev. Mod. Phys.*, **29**:651–663, 1957.
10. K. M. Case and P. F. Zweifel. *Linear Transport Theory*. Addison-Wesley, Reading, 1967.
11. D. G. Collins, W. G. Blaettner, M. B. Wells, and H. G. Horak. Backward Monte Carlo calculations of the polarization characteristics of the radiation emerging from spherical shell atmospheres. *Appl. Opt.*, **11**:2684–2696, 1972.
12. A. Dahlback and K. Stamnes. A new spherical model for computing the radiation field available for photolysis and heating at twilight. *Planet. Space Sci.*, **39**:671–683, 1991.
13. A. Doicu, T. Trautmann, F. Schreier, S. Slijkhuis, S. Hilgers, A. von Bargen, M. Hess, and B. Aberle. Radiative transfer models for a spherical shell atmosphere. *in preparation*.
14. L. Flynn, J. Hornstein, and E. Hilsenrath. The ozone mapping and profiler suite (omps): The next generation of us ozone monitoring instruments. *IEEE IGARSS Proceedings, Anchorage, Alaska*, 2004.

15. O. P. Hasekamp and J. Landgraf. A linearized vector radiative transfer model for atmospheric trace gas retrieval. *J. Quant. Spectrosc. Radiat. Transfer*, **75**:221–238, 2002.
16. O. P. Hasekamp and J. Landgraf. Linearization of vector radiative transfer with respect to aerosol properties and its use in satellite remote sensing. *J. Geophys. Res.*, D04203:doi:10.1029/2004JD005260, 2005.
17. O. P. Hasekamp, J. Landgraf, and R. van Oss. The need of polarization modeling for ozone profile retrieval from backscattered sunlight. *J. Geophys. Res.*, **107**:4692, 2002.
18. B. M. Herman, A. Ben-David, and K. J. Thome. Numerical technique for solving the radiative transfer equation for a spherical shell atmosphere. *Appl. Opt.*, **33**:1760–1770, 1994.
19. J. W. Kaiser and J. P. Burrows. Fast weighting functions for retrievals from limb scattering measurements. *J. Quant. Spectrosc. Radiat. Transfer*, **77**:273–283, 2003.
20. P. Kunasz and L. H. Auer. Short characteristic integration of radiative transfer problems: formal solution in two-dimensional slabs. *J. Quant. Spectrosc. Radiat. Transfer*, **39**:67–79, 1988.
21. J. Landgraf and O. P. Hasekamp. Ozone profile retrieval from satellite measurements of nadir backscattered light in the ultraviolet of the solar spectrum. *Recent Res. Devel. Geophysics*, **4**:157–189, 2002.
22. J. Landgraf, O. P. Hasekamp, and T. Trautmann. Linearization of radiative transfer with respect to surface properties. *J. Quant. Spectrosc. Radiat. Transfer*, **72**:327–339, 2002.
23. J. Landgraf, O. P. Hasekamp, T. Trautmann, and M. A. Box. A linearized radiative transfer model for ozone profile retrieval using the analytical forward-adjoint perturbation theory approach. *J. Geophys. Res.*, **106**:27291–27306, 2001.
24. D. J. Lary and M. Balluch. Solar heating rates: the importance of spherical geometry. *J. Atmos. Sci.*, **50**:3983–3993, 1993.
25. J. Lenoble and Z. Sekera. Equation of radiative transfer in a planetary spherical atmosphere. *Proc. Nat. Acad. Sci.*, **47**:372–389, USA, 1961.
26. J. Lewins. *Importance. The Adjoint Function. The Physical Basis of Variational and Perturbation Theory in Transport and Diffusion Problems*. Pergamon Press, New York, 1965.
27. R. P. Loughman, E. Griffioen, L. Oikarinen, O. V. Postylyakov, A. Rozanov, D. E. Flittner, and D. F. Rault. Comparison of radiative transfer models for limb-viewing scattered sunlight measurements. *J. Geophys. Res.*, **109**:D06303, doi:10.1029/2003JD003854, 2004.
28. G. Marchuk. Equation for the value of information from weather satellites and formulation of inverse problems. *Cosmic Res.*, **2**:394–409, 1964.
29. G. I. Marchuk, G. A. Mikhailov, M. A. Nazaraliev, R. A. Darbinjan, B. A. Kargin, and B. S. Elepov. *The Monte Carlo Methods in Atmospheric Optics*. Springer Verlag, Berlin, 1980.
30. A. Marshak and A. B. Davis. *3D Radiative Transfer in Cloudy Atmospheres*. Springer, New York, 2005.
31. C. A. McLinden, J. C. McConnell, E. Griffioen, and C. T. McElroy. A vector radiative-transfer model for the odin/osiris project. *Can. J. Phys.*, **80**:375–393, 2002.
32. M. I. Mishchenko, A. A. Lacis, and L. D. Travis. Errors induced by the neglect of polarization in radiance calculations for rayleigh-scattering atmospheres. *J. Quant. Spectrosc. Radiat. Transfer*, **51**:491–510, 1994.
33. M. I. Mishchenko, L. D. Travis, and A. A. Lacis. *Scattering, Absorption and Emission of Light by Small Particles*. Cambridge University Press, Cambridge, 2002.

34. National Oceanic and Atmospheric Administration (NOAA). *U.S. Standard Atmosphere*. Rep. NOAA-S/T76-1562, Washington, D.C., U.S. Gov. Print. Off., 1976.
35. L. Oikarinen. Effect of surface albedo variations on uv-visible limb-scattering measurements of the atmosphere. *J. Geophys. Res.*, **107**:doi:10.1029/2001JD001492, 2002.
36. L. Oikarinen, E. Sihvola, and E. Kyrola. Multiple scattering in limb viewing geometry. *J. Geophys. Res.*, **104**:31261–31274, 1999.
37. P. T. Partain, A. K. Heidinger, and G. L. Stephens. High spectral resolution atmospheric radiative transfer: Application of the equivalence theorem. *J. Geophys. Res.*, **105**:2163–2177, 2000.
38. W. F. Payne, E. J. Llewellyn, and J. S. Matsushita. Osiris: An imaging spectrograph for odin. *46th International Astronautical Congress, Oslo, Norway*, 1995.
39. P. Phillips. A technique for the numerical solution of certain integral equations of the first kind. *J. Assoc. Comput. Mach.*, **9**:84–97, 1962.
40. O. V. Postylyakov. Linearized vector radiative transfer model mcc++ for a spherical atmosphere. *J. Quant. Spectrosc. Radiat. Transfer*, **88**:297–317, 2004.
41. C. Rodgers. Retrieval of atmospheric temperature and composition from remote measurements of thermal radiation. *Rev. Geophys.*, **14**:609–624, 1976.
42. A. Rozanov, V. Rozanov, and J. P. Burrows. A numerical radiative transfer model for a spherical planetary atmosphere: combined differential-integral approach involving the picard iterative approximation. *J. Quant. Spectrosc. Radiat. Transfer*, **69**:491–512, 2001.
43. A. Rozanov, V. Rozanov, and J. P. Burrrows. Combined differential-integral approach for the radiation field computation in a spherical shell atmosphere: non-limb geometry. *J. Geophys. Res.*, **105**:22937–43, 2000.
44. V. V. Rozanov, M. Buchwitz, K. U. Eichmann, R. de Beek, and J. P. Burrows. Sciatran – a new radiative transfer model for geophysical applications in the 240–2400 nm spectral region: the pseudo-spherical version. *Adv. Space Res.*, **29**:1831–1835, 2002.
45. V. V. Rozanov, T. Kurosu, and J. P. Burrrows. Retrieval of atmospheric constituents in the uv-visible: A new quasi-analytical approach for the calculation of weighting functions. *J. Quant. Spectrosc. Radiat. Transfer*, **60**:277–99, 1998.
46. K. K. Sen and S. J. Wilson. *Radiative Transfer in Curved Media*. World Scientific, Singapore, 1990.
47. O. I. Smokty. Multiple light scattering in the spherical planetary atmosphere. *Pure Appl. Geophys.*, **72**:214–226, 1969.
48. V. Sobolev. *Light Scattering in Planetary Atmospheres*. Pergamon Press, Oxford, 1975.
49. V. V. Sobolev and I. N. Minin. Light scattering in the spherical atmosphere. *ISZ (artificial earth satellites)*, **14**, 1962.
50. F. Spada, M. C. Krol, and P. Stammes. Mcscia: Application of the equivalence theorem in a monte carlo radiative transfer model for spherical shell atmospheres. *Atmos. Chem. Phys. Discuss.*, **6**:1199–1248, 2006.
51. A. Tikhonov. On the solution of incorrectly stated problems and a method of regularization. *Dokl. Akad. Nauk SSSR*, **151**:501–504, 1963.
52. T. Trautmann and M. A. Box. Green's function computation in radiative transfer theory. In de Groot, R. A., and Nadrchal, J. (Eds), *Physics Computing '92*. World Scientific, Singapore, 1993.
53. E. A. Ustinov. Inverse problem of the photometry of the solar radiation reflected by an optically thick planetary atmosphere. 2. Numerical aspects and requirements on the observation geometry. *Cosmic Res.*, **29**:785–800, 1991.

54. E. A. Ustinov. Inverse problem of the photometry of the solar radiation reflected by an optically thick planetary atmosphere. mathematical framework and weighting functions of linearized inverse problem. *Cosmic Res.*, **29**:519–531, 1991.
55. E. A. Ustinov. Inverse problem of the photometry of the solar radiation reflected by an optically thick planetary atmosphere. 3. remote sensing of minor gaseous constituents and an atmospheric aerosol. *Cosmic Res.*, **30**:170–181, 1992.
56. R. J. van der A. Improved ozone profile retrieval from combined nadir/limb observations of SCIAMACHY. *J. Geophys. Res.*, **106**:14583–94, 2001.
57. H. H. Walter and J. Landgraf. Towards linearization of atmospheric radiative transfer in spherical geometry. *J. Quant. Spectrosc. Radiat. Transfer*, **92**:175–200, 2005.
58. H. H. Walter, J. Landgraf, and O. P. Hasekamp. Linearization of a pseudo-spherical vector radiative transfer model. *J. Quant. Spectrosc. Radiat. Transfer*, **85**:251–283, 2004.
59. H. H. Walter, J. Landgraf, F. Spada, and A. Doicu. Linearization of a radiative transfer model in spherical geometry. *J. Geophys. Res.*, **111**:doi:10.1029/2005JD007014, 2006.
60. Q. Yi, M. A. Box, and T. Trautmann. Higher-order radiative perturbation theory. *J. Quant. Spectrosc. Radiat. Transfer*, **84**:105–114, 2004.

5 Convergence acceleration of radiative transfer equation solution at strongly anisotropic scattering

Vladimir P. Budak, Dmitriy A. Klyuykov, and Sergey V. Korkin

5.1 Introduction

All the numerical methods of the radiative transfer equation (RTE) solution are based on the replacement of scattering integral by the finite sum. The main problem of such scattering integral representation is the presence of singularities in the radiance angular distribution (RAD): they cannot be included in any quadrature formula in essence (Krylov, 2006). To solve this problem various methods are used by different researchers: Wiscombe (1977), Thomas and Stamnes (2002), Rozanov et al. (2005) and others. The physical model of radiative transfer theory is the ray approximation. In this approximation any break in the boundary conditions spreads into the depth of the medium and generates singularities in the RAD. Scattering in the medium gives, along with the singularities, an anisotropic part in the light field. The difficulties in the calculation of the anisotropic part of the solution led to various truncation methods of the scattering anisotropy that are not completely true: in the general 3-D medium geometry the core of the problem is not the scattering, but the presence of singularities in the RAD. Krylov (2006) showed that the most effective and accurate way of including the singularities in the quadrature formula is their analytical, perhaps approximate, representation and an analytical integration – the method of singularities elimination.

For eliminating the anisotropic part of the solution we proposed to use the small-angle approximation in the Goudsmit–Saunderson form. It results in a boundary-value problem for the regular part of the solution that is similar to the initial problem, but with the source function on the right-hand side of the equation and modified boundary conditions. The solution of the obtained equation for the regular part is found by the discrete ordinates method. In the case of the plane-parallel geometry this problem has the analytical solution as a matrix exponential. The final solution in this case has the form of the transfer matrix through a slab with a column of sources. On the basis of the matrix-operator method it allows us to proceed to the solution of the problem for the case of a plane-parallel slab composed of several homogeneous layers.

To use the proposed method in the arbitrary 3-D medium geometry we should generalize the Goudsmit–Saunderson approximation. We proposed this approach in our previous work and called it the small-angle modification of the spherical

harmonics method (MSH). From the analysis of the angular spectrum of the radiance distribution, it is possible to formulate the approximate equation describing an anisotropic part of the solution: in the neighborhood of the singularity the spectrum slowly decreases with the number of harmonics. MSH includes all the singularities of the exact solution of the radiative transfer equation. This approach makes it possible to eliminate analytically the singularities from the RTE solution and to state the boundary-value problem for the regular part of the solution. The regular part of the solution is found numerically by the finite element method, which enables us to extend this approach to the arbitrary medium geometry.

For the generalization of this approach in the case of the vectorial RTE (VRTE) for polarized light transfer we used the polarization determination in the basis of the circular polarization (CP-presentation, (Kuščer and Ribarič, 1959). Unfortunately in CP-presentation all the coefficients in VRTE become complex and that creates difficulties for the use of effective numerical solution methods. Therefore for the VRTE solution we follow the following steps:

- convert equation to CP-presentation,
- subtract vectorial MSH,
- obtain the equation for the smooth part,
- apply the addition theorem for generalized spherical harmonic,
- return to SP-presentation,
- solve the obtained equation.

5.2 Singularities of the solution of the radiative transfer equation

The physical model of the radiative transfer theory is the light field theory based on the ray approximation of light radiation (Dolin, 1964; Apresyan and Kravtsov, 1996; Mishchenko, 2008). In optics, the ray approximation is inapplicable for the description of processes at scales smaller than the wavelength of light (Born and Wolf, 1975). On the other hand, ray approximation describes the light propagation accurately enough except for some special cases.

Ray optics is a rigorous closed theory; it can be developed on a phenomenological basis as the theory of the light field of the responses of an optical detector placed at a certain point of this field. The optical detector responds to the square magnitude of the electromagnetic field and averages values in time and space that substantially exceed the period and length of the light wave. The light field consists of independent rays penetrating each point of space and transferring light energy. The measure of the power density in the ray bundle propagating in the direction $\hat{\boldsymbol{l}}$ from the point $\mathbf{r}$ is the radiance of the light field $L(\boldsymbol{r}, \hat{\boldsymbol{l}})$. (Hereinafter we mark the unit vector with symbol: $\hat{}$).

At the transition from wave optics to ray optics, Maxwell equations are converted to the radiative transfer equation (Apresyan and Kravtsov, 1996; Mishchenko, 2008; Lommel, 1889; Chwolson, 1889) describing the radiance of the light field $L(\boldsymbol{r}, \hat{\boldsymbol{l}})$:

$$(\hat{\boldsymbol{l}}, \nabla) L(\boldsymbol{r}, \hat{\boldsymbol{l}}) = -k_{ext} L(\boldsymbol{r}, \hat{\boldsymbol{l}}) + \frac{\omega_0 k_{ext}}{4\pi} \oint p(\hat{\boldsymbol{l}}, \hat{\boldsymbol{l}}') L(\boldsymbol{r}, \hat{\boldsymbol{l}}')\, d\hat{\boldsymbol{l}}' + q(\boldsymbol{r}, \hat{\boldsymbol{l}})\,, \qquad (5.1)$$

where k_{ext} is the attenuation coefficient, $k_{ext} = k_{abs} + k_{sca}$; k_{abs}, k_{sca} are the absorption and the scattering coefficients accordingly; $\omega_0 = k_{sca}/k_{ext}$ is the single scattering albedo; $p(\hat{\boldsymbol{l}}, \hat{\boldsymbol{l}}')$ is the phase function of the elementary volume of a scattering medium; $q(\boldsymbol{r}, \hat{\boldsymbol{l}})$ is the source function of the intrinsic emission of the elementary medium volume. The integration in equation (5.1) is by the full solid angle 4π, and $d\hat{\boldsymbol{l}}'$ is an elementary solid angle in the direction $\hat{\boldsymbol{l}}'$.

To have a single solution, RTE (5.1) must have the boundary conditions. Let us analyze the field inside the volume V, and the radiation $L_e(\boldsymbol{r}, \hat{\boldsymbol{l}})$ is incident on the outer boundary Σ of V (Fig.1). Then the boundary conditions are:

$$L(\boldsymbol{r}, \hat{\boldsymbol{l}})\Big|_{\boldsymbol{r}\in\Sigma,\ (\hat{\boldsymbol{N}},\hat{\boldsymbol{l}})\geq 0} = L_e(\boldsymbol{r}, \hat{\boldsymbol{l}})\,, \tag{5.2}$$

where $\hat{\boldsymbol{N}}$ is the inner normal to the surface Σ in the point $\mathbf{r}$ (Fig. 5.1).

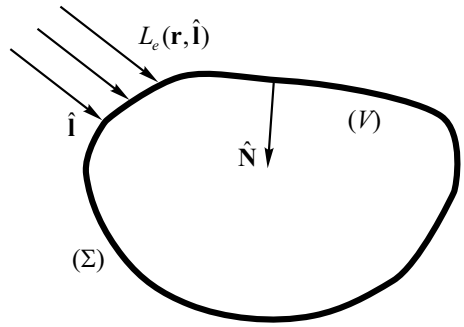


Fig. 5.1. The boundary conditions for the radiative transfer equation.

The distinctive feature of the ray approximation is the following: if there is a break (singularity) in the radiance spatial-angular distribution $L_e(\boldsymbol{r}, \hat{\boldsymbol{l}})$ of the boundary conditions (5.2) it spreads into the medium depth and has a place for the distribution $L(\boldsymbol{r}, \hat{\boldsymbol{l}})$. The most striking example of this phenomenon is the abrupt radiance junction at the border of light and shade. Note that in wave optics, there are no angular singularities due to diffraction.

Since practically all light sources have limited spatial-angular light distribution, there are always singularities in the RTE solution. The most effective solution method of the boundary problems of transport theory is the Green function method (Case et al., 1967), where the solution of (5.1) with conditions (5.2) is presented in the form of the superposition of the solution for the elementary sources

$$\begin{aligned} L(\boldsymbol{r}, \hat{\boldsymbol{l}}) \;=\; & \int\limits_{(V)} \oint G(\boldsymbol{r}_0, \hat{\boldsymbol{l}}_0 \to \boldsymbol{r}, \hat{\boldsymbol{l}}) q(\boldsymbol{r}_0, \hat{\boldsymbol{l}}_0)\, d\hat{\boldsymbol{l}}_0\, d^3 r_0 \\ & + \int\limits_{(\Sigma)} \oint G_\Sigma(\boldsymbol{r}_\Sigma, \hat{\boldsymbol{l}}_\Sigma \to \boldsymbol{r}, \hat{\boldsymbol{l}}) L_e(\boldsymbol{r}_\Sigma, \hat{\boldsymbol{l}}_\Sigma)\, d\hat{\boldsymbol{l}}_\Sigma\, d^2 r_\Sigma\,, \end{aligned} \tag{5.3}$$

where G, G_Σ are the volume and surface Green functions connected with each other (Case and Zweifel, 1967).

All the real problems are reduced to some elementary sources: point unidirectional, point diffuse, point isotropic and plane unidirectional (Germogenova, 1986). Because the elementary sources have bounded light distribution their radiance spatial-angular distribution contains the singularities that are shown in Table 5.1 (Germogenova, 1986). The presence of solution singularities causes some difficulties to solve (5.1). Indeed, any numerical solution method is based on the replacement of the scattering integral in (5.1) by the finite sum. This cannot be done with singularities since the singularities for their presentation in any basis require an infinite number of elements.

Table 5.1. Singularities of RTE solution for elementary source (Germogenova, 1986)

Order of scattering	**Point unidirectional**	**Point isotropic**	**Plane unidirectional**
0	$\dfrac{\delta(\hat{\omega}-\hat{\boldsymbol{l}}_0)\delta(\hat{\boldsymbol{l}}-\hat{\boldsymbol{l}}_0)}{\lvert\boldsymbol{r}-\boldsymbol{r}_0\rvert^2}$	$\dfrac{\delta(\hat{\omega}-\hat{\boldsymbol{l}}_0)}{\lvert\boldsymbol{r}-\boldsymbol{r}_0\rvert^2}$	$\dfrac{\delta(\hat{\boldsymbol{l}}-\hat{\boldsymbol{l}}_0)}{\lvert\mu_0\rvert}\mathrm{e}^{-\frac{k_{ext}z}{\mu_0}}$
1	$\dfrac{\delta(\psi-\eta)}{\rho\sqrt{1-(\hat{\omega},\hat{\boldsymbol{l}}_0)^2}}$	$\dfrac{1}{\lvert\boldsymbol{r}-\boldsymbol{r}_0\rvert\sqrt{1-(\hat{\omega},\hat{\boldsymbol{l}}_0)^2}}$	
2	$\dfrac{\ln(\rho\lvert\psi-\eta\rvert s)}{\sqrt{1-(\hat{\omega},\hat{\boldsymbol{l}}_0)^2}}$	$\ln\left(\lvert\boldsymbol{r}-\boldsymbol{r}_0\rvert(1-(\hat{\omega},\hat{\boldsymbol{l}}_0)^2)\right)$	
3	$\dfrac{1}{\sqrt{1-(\hat{\omega},\hat{\boldsymbol{l}}_0)^2}}$		

where $\mu=(\hat{\boldsymbol{l}},\hat{\boldsymbol{z}})$, $\hat{\boldsymbol{l}}=\{\mu\cdot\cos\psi,\ \mu\cdot\sin\psi,\ \mu\}$, ψ is the azimuth in the XOY-plane;

$$\boldsymbol{r}=\{\rho,z,\eta\},\ \hat{\omega}=\frac{\boldsymbol{r}-\boldsymbol{r}_0}{\lvert\boldsymbol{r}-\boldsymbol{r}_0\rvert},\ s=\sqrt{1-(\hat{\omega},\hat{\boldsymbol{l}})}.$$

One of the boundary problems of RTE is the light field in the scattering medium slab with the depth d irradiated by plane unidirectional sources at an angle θ_0 to the normal of the slab border. It corresponds to the problem of the light field determination in the atmosphere or in the ocean when the Sun is seen from the Earth at small angle size ($\sim 32'$). In this case, if we introduce the Cartesian coordinate system $OXYZ$ with the axis OZ along inner normal to the upper slab border:

$$\begin{aligned}
&\boldsymbol{r}=x\hat{\boldsymbol{x}}+y\hat{\boldsymbol{y}}+z\hat{\boldsymbol{z}},\ \mu=(\hat{\boldsymbol{l}},\hat{\boldsymbol{z}}),\\
&\hat{\boldsymbol{l}}=\{\sqrt{1-\mu^2}\cos\varphi,\ \sqrt{1-\mu^2}\sin\varphi,\ \mu\},\\
&L\equiv L(z,\mu,\varphi):\ (\hat{\boldsymbol{l}},\nabla)L=\mu\frac{\partial L}{\partial z},
\end{aligned}\tag{5.4}$$

that reduces the problem solution to the boundary one

$$\begin{cases} \mu\dfrac{\partial L}{\partial\tau} + L(\tau,\mu,\varphi) = \dfrac{\omega_0}{4\pi}\oint p(\hat{\boldsymbol{l}}',\hat{\boldsymbol{l}})L(\tau,\mu',\varphi')\,d\hat{\boldsymbol{l}}', \\ L(\tau,\mu,\varphi)|_{\tau=0,\ \mu>0} = \delta(\hat{\boldsymbol{l}}-\hat{\boldsymbol{l}}_0), \quad L(\tau,\mu,\varphi)|_{\tau=\tau_0,\ \mu<0} = 0; \end{cases} \tag{5.5}$$

where $\tau = \int\limits_0^z k_{ext}(z')\,dz'$ is the optical depth, $\tau_0 = \int\limits_0^d k_{ext}(z')\,dz'$ is the optical thickness of slab, $\hat{\boldsymbol{l}}_0 = \left\{\sqrt{1-\mu_0^2}\cos\varphi_0,\ \sqrt{1-\mu_0^2}\sin\varphi_0,\ \mu_0\right\}$, $\mu_0 = \cos\theta_0$. Everywhere else, except the special cases, we consider the slab as homogeneous ($\tau = k_{ext}z$) and for simplicity assume that $\varphi_0 = 0$.

It is seen from the boundary conditions (5.5) that they have δ-singularity in the angular distribution. To eliminate them, Milne (1926) proposed to subtract the direct nonscattered radiation from the full radiance of the light field

$$L(\tau,\mu,\varphi) = \mathrm{e}^{-\tau/\mu_0}\delta(\hat{\boldsymbol{l}}-\hat{\boldsymbol{l}}_0) + L_d(\tau,\mu,\varphi) \tag{5.6}$$

and to formulate the boundary problem only for the diffuse light field component:

$$\begin{cases} \mu\dfrac{\partial L_d}{\partial\tau} + L_d(\tau,\mu,\varphi) = \dfrac{\omega_0}{4\pi}\oint p(\hat{\boldsymbol{l}}',\hat{\boldsymbol{l}})L_d(\tau,\mu',\varphi')\,d\hat{\boldsymbol{l}}' + \dfrac{\omega_0}{4\pi}\,\mathrm{e}^{-\tau/\mu_0}p(\hat{\boldsymbol{l}}_0,\hat{\boldsymbol{l}}), \\ L_d(\tau,\mu,\varphi)|_{\tau=0,\ \mu>0} = 0, \quad L_d(\tau,\mu,\varphi)|_{\tau=\tau_0,\ \mu<0} = 0. \end{cases} \tag{5.7}$$

This approach was developed by Chandrasekhar (1950) who applied them for the solution of different problems with isotropic and Rayleigh phase functions. However, all the scattering media, whether in the atmosphere or in the ocean, have suspended particles of sizes much greater than the wavelength. According to Mie theory this results in strong anisotropic light scattering. The application of this approach to anisotropic scattering encounters the serious difficulties that the different procedures of smoothing or truncation of the phase function (Thomas and Stamnes, 2002) tried to overcome.

It is not difficult to understand that the scattering integral is the convolution by the solid angle on the sphere of the two functions: the phase function $p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}')$ and the radiance distribution $\tilde{L}(\boldsymbol{r},\hat{\boldsymbol{l}}')$. The calculation accuracy depends on the accuracy of presentation of both functions. If we turn to Table 5.1, we see that the RAD for the elementary sources contains the angular singularities, not only in the direct radiation but also in the first two orders of scattering.

The optimal method to remove the anisotropy problem in the RAD is the approximate presentation of the anisotropic part of the solution $L_a(\tau,\mu,\varphi)$ and the presentation of the full solution as the sum:

$$L(\tau,\mu,\varphi) = L_a(\tau,\mu,\varphi) + \tilde{L}(\tau,\mu,\varphi), \tag{5.8}$$

This results to the boundary problem

$$\begin{cases} \mu\dfrac{\partial\tilde{L}}{\partial\tau} + \tilde{L}(\tau,\mu,\varphi) = \dfrac{\omega_0}{4\pi}\oint p(\hat{\boldsymbol{l}}',\hat{\boldsymbol{l}})\tilde{L}(\tau,\mu',\varphi')\,d\hat{\boldsymbol{l}}' + S(\tau,\mu,\varphi), \\ \tilde{L}(\tau,\mu,\varphi)\Big|_{\tau=0,\ \mu>0} = 0, \quad \tilde{L}(\tau,\mu,\varphi)\Big|_{\tau=\tau_0,\ \mu<0} = -L_a(\tau_0,\mu,\varphi), \end{cases} \tag{5.9}$$

where the source function (RTE discrepancy that is due to an inaccuracy of the anisotropy part) is:

$$S(\tau,\mu,\varphi)=\frac{\omega_0}{4\pi}\oint p(\hat{\boldsymbol{l}}',\hat{\boldsymbol{l}})L_a(\tau,\mu',\varphi')\,d\hat{\boldsymbol{l}}'-\mu\frac{\partial L_a}{\partial\tau}-L_a(\tau,\mu,\varphi). \tag{5.10}$$

The modification of the boundary conditions (as it will be shown below) was necessary because the forward radiation is taken into account in the anisotropic part very strictly while the backscattering radiation is taken approximately. So, this inaccuracy has to be compensated to make the boundary conditions (5.7) and (5.9) equivalent.

Note that the requirements of the scattering integral presentation accuracy are determined by the most slowly varying function. Therefore, if the anisotropic part is extracted from the solution accurately enough, then the requirements of the scattering integral presentation will be determined by the smooth regular part, irrespective of the scattering anisotropy.

5.3 Small angle modification of the spherical harmonics method

To extract the anisotropic part from the solution, one has to state its distinguishing feature. It could be much easier, if we come over to the spectral representation of the angular distribution, i.e. to the angular spectrum of the radiance distribution. For the function given on the sphere (zenith θ and azimuth φ angles), the spectral representation is the expansion in series on surface harmonics (Vilenkin, 1965), that is used at the RTE solution by the spherical harmonics (SH) method. This method was used by Lorentz (1916) for electron transfer theory. This method was modified for the RTE solution in the discussion between Jeans (1917) and Eddington (1916), and its final formulation in a modern form was proposed by Gratton (1939). The anisotropic part of the solution at the small optical depth has to be approximately symmetrical with respect to the incident direction $\hat{\boldsymbol{l}}_0$. Therefore, in contrast to the traditional presentation, we expand the RAD in series with respect to $\hat{\boldsymbol{l}}_0$:

$$\begin{aligned}L_a(\tau,\mu,\varphi) &= \sum_{k=0}^{\infty}\sum_{m=-\infty}^{\infty}\frac{2k+1}{4\pi}\sqrt{\frac{(k-m)!}{(k+m)!}}C_k^m(\tau)P_k^m(\nu)\,\mathrm{e}^{im\varphi}\\ &= \sum_{k=0}^{\infty}\sum_{m=-\infty}^{\infty}\frac{2k+1}{4\pi}C_k^m(\tau)Q_k^m(\nu)\,\mathrm{e}^{im\varphi}\,.\end{aligned} \tag{5.11}$$

where $\nu=(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0)$, $\mu=\nu\mu_0+\sqrt{1-\nu^2}\,\sqrt{1-\mu_0^2}\,\cos\varphi$, $P_l^n(\mu)$ is the associated Legendre polynomial, and $P_l^0(\mu)\equiv P_l(\mu)$ is the Legendre polynomial;

$$Q_l^n(\mu)=\sqrt{\frac{(l-n)!}{(l+n)!}}P_l^n(\mu)$$

is the renormalized Legendre polynomial.

We substitute the presentation (5.8), (5.11) in RTE (5.5), multiply it by $Q_k^m(\nu)\,\mathrm{e}^{im\varphi}$ and integrate it over the complete solid angle. Taking into account the orthogonality, the recurrence relationships, and the addition theorems for the Legendre polynomials (Vilenkin, 1965) we get the coupled differential equations

$$\begin{aligned}
&\mu_0 \frac{d}{d\tau}\left[\sqrt{(k-m)(k+m)}C_{k-1}^m(\tau) + \sqrt{(k+m+1)(k-m+1)}C_{k+1}^m(\tau)\right] \\
&\quad + \frac{\sqrt{1-\mu_0^2}}{2}\frac{d}{d\tau}\left\{\left[\sqrt{(k-m+1)(k-m+2)}C_{k+1}^{m-1}(\tau)\right.\right. \\
&\qquad\qquad \left. - \sqrt{(k+m)(k+m-1)}C_{k-1}^{m-1}(\tau)\right] \\
&\qquad\qquad + \left[\sqrt{(k+m+1)(k+m+2)}C_{k+1}^{m+1}(\tau)\right. \\
&\qquad\qquad \left.\left. - \sqrt{(k-m-1)(k-m)}C_{k-1}^{m+1}(\tau)\right]\right\} \\
&= -(2k+1)(1-\omega_0 p_k)C_k^m(\tau)\,, \qquad (5.12)
\end{aligned}$$

where p_k are the expansion coefficients of the phase function on the Legendre polynomials:

$$p(\hat{\boldsymbol{l}}, \hat{\boldsymbol{l}}') = \sum_{k=0}^{N_x} (2k+1) p_k P_k(\hat{\boldsymbol{l}} \cdot \hat{\boldsymbol{l}}')\,, \qquad (5.13)$$

N_x is the series term number in the expansion of the phase function on the Legendre polynomials. We assume for simplicity that $p_k \equiv 0$ at $k > N_x$.

The developed system is different from the conventional one used in the SH method for plane unidirectional source (Budak and Sarmin, 1990) by the different spherical coordinate system in the direction space of radiance sighting with respect to an axis $\hat{\boldsymbol{l}}_0$ instead of $\hat{\boldsymbol{z}}$.

Since the radiance of the anisotropic part of the solution part changes quickly with the angle of sight, its spectrum $C_k^m(\tau)$ has conversely a slow dependence on the index k. It allows us to make the following assumptions.

Continuous dependence of the expansion coefficients (5.11) on the harmonic number $C_k^m(\tau) = C^m(\tau, k)$, which is a slowly monotonous function owing to the angular anisotropy that makes it possible to present it as

$$C^m(\tau, k \pm 1) \approx C^m(\tau, k) \pm \frac{\partial C^m(\tau, k)}{\partial k}\,; \qquad (5.14)$$

Terms with the numbers $k \gg 1$ makes the major contribution to the series (5.11), and the degree of anisotropy is substantially larger than its azimuthal asymmetry ($k \gg m$) that allows us to make the approximations:

$$\begin{aligned}
&\sqrt{(k-m)(k+m)} \approx k; \\
&\sqrt{(k+m+1)(k-m+1)} \approx k+1; \\
&\sqrt{(k-m+1)(k-m+2)} \approx \kappa - \tfrac{(m-1)(2k+1)}{2\kappa}; \\
&\sqrt{(k+m)(k+m-1)} \approx \kappa + \tfrac{(m-1)(2k+1)}{2\kappa}; \\
&\sqrt{(k+m+1)(k+m+2)} \approx \kappa + \tfrac{(m+1)(2k+1)}{2\kappa}; \\
&\sqrt{(k-m)(k-m-1)} \approx \kappa - \tfrac{(m+1)(2k+1)}{2\kappa};
\end{aligned}$$

where $\kappa = \sqrt{k(k+1)}$.

By substituting these assumptions in (5.12) and by replacing k with κ we get:

$$\mu_0 \frac{\partial C^m}{\partial \tau} + \frac{\sqrt{1-\mu_0^2}}{2} \frac{\partial}{\partial \tau} \left[\frac{\partial C^{m+1}}{\partial k} + \frac{\partial C^{m-1}}{\partial k} + \frac{1}{\kappa} \left((m+1)C^{m+1} - (m-1)C^{m-1} \right) \right] = -(1-\omega_0 p_k) C^m(\tau, \kappa)\,. \tag{5.15}$$

In equation (5.15) the term with the second derivation is removed taking into account the condition $k \gg 1$ and the assumption of the slowly monotonic dependence C^m on the number k.

We introduce a function $f(\tau, \boldsymbol{l}_{0\perp}, \boldsymbol{\kappa})$ with its azimuthal spectrum equal to the required function

$$f(\tau, \boldsymbol{l}_{0\perp}, \kappa) = \sum_{m=-\infty}^{\infty} C^m(\tau, \kappa)\, \mathrm{e}^{im\psi}, \tag{5.16}$$

where

$$C^m(\tau, \kappa) = \frac{1}{2\pi} \int_0^{2\pi} f(\tau, \boldsymbol{l}_{0\perp}, \kappa)\, \mathrm{e}^{-im\psi}\, d\psi\,.$$

It is not difficult to derive the following properties of the series (5.16):

$$\sum_{m=-\infty}^{\infty} C^{m\pm1}(\tau, \kappa)\, \mathrm{e}^{im\psi} = \mathrm{e}^{\mp i\psi} f(\tau, \boldsymbol{l}_{0\perp}, \boldsymbol{\kappa}),$$
$$\sum_{m=-\infty}^{\infty} (m \pm 1) C^{m\pm1}(\tau, \kappa)\, \mathrm{e}^{im\psi} = -i\, \mathrm{e}^{\mp i\psi} \frac{\partial f(\tau, \boldsymbol{l}_{0\perp}, \boldsymbol{\kappa})}{\psi}. \tag{5.17}$$

We multiply equation (5.15) by $\mathrm{e}^{im\psi}$ and sum it over m from $-\infty$ to $+\infty$. Having in mind the properties (5.17) we get the followings expression

$$\sum_{m=-\infty}^{\infty} \left[\frac{\partial C^{m+1}}{\partial \kappa} + \frac{\partial C^{m-1}}{\partial \kappa} + \frac{1}{\kappa} \left((m+1)C^m - (m-1)C^m \right) \right] \mathrm{e}^{im\psi} = 2 \left(\cos\psi \frac{\partial f}{\partial \kappa} - \frac{\sin\psi}{\kappa} \frac{\partial f}{\partial \psi} \right) = 2(\hat{\boldsymbol{l}}_{0\perp}, \nabla_\kappa) f\,,$$

that changes equation (5.15) to

$$\frac{\partial}{\partial \tau} \left[\mu_0 f + (\boldsymbol{l}_{0\perp}, \nabla_\kappa) f \right] = -(1-\omega_0 p_k) f(\tau, \boldsymbol{l}_{0\perp}, \kappa)\,. \tag{5.18}$$

The result obtained has a complicated analytical solution. It is important for us to achieve a simple analytical solution with the formulation of the boundary problem for the regular part of (5.9). The item $(\boldsymbol{l}_{0\perp}, \nabla_\kappa) f$ determines the relationship between the harmonic amplitudes $C_k^m(\tau)$ and different azimuthal numbers m. This allows us to describe the effect of the maximum RAD rotation from the incident light direction to the vertical direction in deep layers of the medium.

This phenomenon could be ignored when solving the anisotropic part, since the angular distribution ceases to be anisotropic in the light depth mode. Neglecting

this item we get the following equation:

$$\mu_0 \frac{\partial}{\partial \tau} C_k^m(\tau) = -(1 - \omega_0 p_k) C_k^m(\tau)\,, \tag{5.19}$$

Eliminating the item $(\boldsymbol{l}_{0\perp}, \nabla_\kappa)$ in equation (5.18) means that we neglect the path variance of the scattering rays and the backscattering in the solution. Thus we have the boundary conditions for equation (5.19):

$$C_k^m(0) = 1\,.$$

Taking into account this boundary condition the solution of equation (5.19) has a simple analytical expression

$$C_k^m(\tau) = \exp\left(-(1 - \omega_0 p_k)\tau/\mu_0\right) \equiv Z_k(\tau)\,, \tag{5.20}$$

or for the anisotropic part of the radiance angular distribution:

$$L_a(\tau, \mu, \varphi) = \sum_{k=0}^{\infty} \sum_{m=-\infty}^{\infty} \frac{2k+1}{4\pi} Z_k(\tau) Q_k^m(\nu)\, \mathrm{e}^{im\varphi}\,. \tag{5.21}$$

It is easy to see from equation (5.20) that RAD has an axial symmetry and all the angular coefficients do not depend on the azimuthal index m.

For further analysis it is more convenient to return in (5.21) to the coordinate system (μ,φ) relative to axis OZ that is orthogonal to the medium border since we will find the regular supplement $\tilde{L}(\tau, \mu, \varphi)$ in this coordinate system:

$$L_a(\tau, \mu, \varphi) = \sum_{l=0}^{\infty} \sum_{n=-l}^{l} \frac{2l+1}{4\pi} Z_l^n(\tau) Q_l^n(\mu)\, \mathrm{e}^{in\varphi}\,. \tag{5.22}$$

The expression (5.22) has no symmetry relative to axis OZ and in this case the coefficients $Z_l^n(\tau)$ depend on the azimuthal index n.

Multiply (5.21) and (5.22) by $Q_l^n(\mu)\, \mathrm{e}^{-in\varphi}$ and integrate them over entire solid angle. Since $Q_l^n(\mu)\, \mathrm{e}^{in\varphi}$ is orthogonal we get from the equality of (5.21) and (5.22):

$$Z_l^n(\tau) = \sum_{k=0}^{\infty} \sum_{m=-\infty}^{\infty} \frac{2k+1}{4\pi} Z_k(\tau) \oint Q_k^m(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}_0)\, \mathrm{e}^{im\varphi} Q_l^n(\hat{\boldsymbol{l}}\hat{\boldsymbol{z}})\, \mathrm{e}^{-in\varphi}\, d\hat{\boldsymbol{l}}\,. \tag{5.23}$$

It is not difficult to compute the last expression on the basis of the addition theorem for the Legendre polynomials:

$$Z_k^m(\tau) = Z_k(\tau) Q_k^m(\mu_0)\,. \tag{5.24}$$

Below we call the obtained algorithm of the determination of the anisotropic part of the solution as the small angle modification of the spherical harmonics method, in short: the modification of the spherical harmonics method (MSH) (Budak and Sarmin, 1990; Budak and Korkin, 2008a). The solution obtained is in good agreement with the small angle approximation (SA) in the Goudsmit–Saunderson form (Goudsmit and Saunderson, 1940).

5.4 Small angle approximation in transport theory

SA appeared for the first time in Wentzel (1922) in the analysis of electron multiple scattering in a solid body. The analysis was carried out on the basis of the expression for the arbitrary order of scatterings, which can be written in our designation for the radiance of the n-order scattering light in the form:

$$L_n(\tau,\hat{\boldsymbol{l}}) = \left(\frac{\omega_0}{4\pi}\right)^n \underbrace{\int_0^{\xi_n}\cdots\int_0^{\xi_1}}_{n} \underbrace{\oint\cdots\oint}_{n} \exp\left(-\xi-\sum_{k=1}^{n}\zeta_k\right) \cdot p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_n)\cdot\ldots\cdot p(\hat{\boldsymbol{l}}_2,\hat{\boldsymbol{l}}_1)L_0(\hat{\boldsymbol{l}}_1)\,d\hat{\boldsymbol{l}}_1\cdots d\hat{\boldsymbol{l}}_n\,d\zeta_1\cdots d\zeta_n\,. \tag{5.25}$$

The ray trajectory at the n-order scattering is a polyline from the $n+1$ segments. The optical depth and the direction of the first n segments from plane unidirectional source are designated by ζ_i, $\hat{\boldsymbol{l}}_i$ respectively. The last segment adjacent to the detector has a length $\xi = \tau/\mu - \sum_{k=1}^{n}\zeta_k$. Since the electron scattering has a strongly anisotropic character, the length variance of the scattering ray path for small angles can be neglected (Wentzel, 1922)

$$\xi + \sum_{k=1}^{n}\zeta_k \approx \frac{\tau}{\mu_0}\,, \tag{5.26}$$

With substitution of (5.26) the equation (5.25) is transformed to

$$L_n(\tau,\hat{\boldsymbol{l}}) = \frac{1}{n!}\left(\frac{\omega_0\tau}{4\pi\mu_0}\right)^n \exp\left(-\frac{\tau}{\mu_0}\right) \underbrace{\oint\cdots\oint}_{n} p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_1)\cdot\ldots\cdot p(\hat{\boldsymbol{l}}_{n-1},\hat{\boldsymbol{l}}_n)L_0(\hat{\boldsymbol{l}}_n)\,d\hat{\boldsymbol{l}}_1\cdots d\hat{\boldsymbol{l}}_n\,.$$

Thus, the radiance of a multiply scattered light can be described as:

$$\begin{aligned} L(\tau,\hat{\boldsymbol{l}}) &= \sum_{n=0}^{\infty} L_n(\tau,\hat{\boldsymbol{l}}) \\ &= \mathrm{e}^{-\tau/\mu_0}\sum_{n=0}^{\infty}\frac{(\omega_0\tau/\mu_0)^n}{n!}\left(\frac{1}{4\pi}\right)^n \underbrace{\oint\cdots\oint}_{n} p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_n)\cdot\ldots\cdot p(\hat{\boldsymbol{l}}_2,\hat{\boldsymbol{l}}_1)\,d\hat{\boldsymbol{l}}_1\cdots d\hat{\boldsymbol{l}}_n \\ &= \mathrm{e}^{-\tau/\mu_0}\sum_{n=0}^{\infty}\frac{(\omega_0\tau/\mu_0)^n}{n!}\Phi_n(\hat{\boldsymbol{l}})\,. \end{aligned} \tag{5.27}$$

It seems to be obvious that small angle approximation is equivalent to the RTE (5.1) transformation to

$$(\hat{\boldsymbol{l}}_0,\nabla)L(\tau,\hat{\boldsymbol{l}}) = -L(z,\hat{\boldsymbol{l}}) + \frac{\omega_0}{4\pi}\oint p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}')L(\tau,\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}'\,. \tag{5.28}$$

In this case the equation for the plane unidirectional source has the following form

$$\mu_0 \frac{\partial L}{\partial \tau} = -L(\tau, \mu) + \frac{\omega_0}{4\pi} \oint p(\hat{\boldsymbol{l}}, \hat{\boldsymbol{l}}') L(\tau, \mu')\, d\hat{\boldsymbol{l}}' \,, \tag{5.29}$$

or that is equivalent in the integral form to the next equation (for $\mu > 0$):

$$L(\tau, \mu) = \delta(\hat{\boldsymbol{l}} - \hat{\boldsymbol{l}}_0)\, \mathrm{e}^{-\tau/\mu_0} + \frac{\omega_0}{4\pi\mu_0} \int_0^\tau \mathrm{e}^{-\xi/\mu_0} \oint p(\hat{\boldsymbol{l}}, \hat{\boldsymbol{l}}') L(\tau, \mu')\, d\hat{\boldsymbol{l}}' d\xi \,. \tag{5.30}$$

Now we can ascertain that the solution of equation (5.30) coincides with the expression (5.27) by the straightforward calculation.

It could be seen from (5.27) that the central problem in SA is the computation of the multi-dimensional convolution over solid angle on the sphere $\Phi_n(\hat{\boldsymbol{l}})$. There are three ways of solving the problem.

In the first approach (Bothe, 1929), the field radiance in the direction $\hat{\boldsymbol{l}}'$ was presented in the expanded form of a Taylor series

$$L(z, \boldsymbol{r}, \hat{\boldsymbol{l}}') = \sum_{n=0}^{\infty} \frac{(\hat{\boldsymbol{l}}' - \hat{\boldsymbol{l}})^n}{n!} \nabla_{\boldsymbol{l}}^n L \,.$$

These series are substituted in the scattering integral of RTE (5.28) where the first three terms are preserved and with the assumption that the phase function is sharper angle function than the radiance distribution. As the result, the RTE (5.28) is reduced to the equation of Fokker–Planck type (diffuse approximation):

$$(\hat{\boldsymbol{l}}_0, \nabla) L(z, \hat{\boldsymbol{l}}) + \kappa L(z, \hat{\boldsymbol{l}}) = \frac{\langle \vartheta^2 \rangle}{2} \Delta_{\boldsymbol{l}} L, \tag{5.31}$$

where

$$\Delta_{\boldsymbol{l}} = \frac{\partial^2}{\partial l_x^2} + \frac{\partial^2}{\partial l_y^2}$$

is Laplacian in the space of the radiation sighting direction and l_x, l_y are the projections of the vector $\hat{\boldsymbol{l}}$ on the plane perpendicular to the direction $\hat{\boldsymbol{l}}_0$.

The solution of (5.29) is just the gaussoid on the sighting angle. Here $\nabla_{\boldsymbol{l}}^n$ is the gradient in the space of the radiation sighting direction, and

$$\langle \vartheta^2 \rangle = \frac{1}{4\pi} \oint p(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')(\hat{\boldsymbol{l}} - \hat{\boldsymbol{l}}')^2\, d\hat{\boldsymbol{l}}' \approx \frac{1}{2} \int_0^\pi p(\cos\vartheta)\vartheta^2 \sin\vartheta d\vartheta \,.$$

The additional assumption about the greater anisotropic degree of the phase function in comparison with the RAD smooths strongly all the solution singularities. But this is valid only for the small angles range in deep layers of a scattering medium.

Another form of SA based on the addition theorem for Legendre polynomials was presented by Goudsmit and Saunderson, (1940). If $L_0(\boldsymbol{r}, \hat{\boldsymbol{l}})$ and $x(\hat{\boldsymbol{l}} \cdot \hat{\boldsymbol{l}}')$ are presented as the series on the surface functions similar to (5.13) then the expression

of convolution for plane unidirectional source is

$$\Phi_n(\hat{\boldsymbol{l}}) = \sum_{k=0}^{\infty} \frac{2k+1}{4\pi} p_k^n P_k(\hat{\boldsymbol{l}} \cdot \hat{\boldsymbol{l}}_0)\,. \tag{5.32}$$

It follows from (5.30) that

$$L(\tau, \hat{\boldsymbol{l}}) = \mathrm{e}^{-\tau/\mu_0} \sum_{k=0}^{\infty} \frac{2k+1}{4\pi} \exp\left(\frac{\omega_0 \tau p_k}{\mu_0}\right) P_k(\hat{\boldsymbol{l}} \cdot \hat{\boldsymbol{l}}_0)\,. \tag{5.33}$$

The third SA form was presented a long time ago (Kompaneets, 1947; Molière, 1948; Snyder and Scott, 1949). In those works, along with the neglect of the ray path variance, it was assumed that in the range of small angles, convolution over the sphere (rotation) can be replaced by convolution over the plane tangent to the sphere in $\hat{\boldsymbol{l}}_0$ (plane translation):

$$\Phi_n(\hat{\boldsymbol{l}}) \approx \underbrace{\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty}}_{n} p(\boldsymbol{l}_\perp - \boldsymbol{l}_{\perp n}) \cdot \ldots \cdot p(\boldsymbol{l}_{\perp 2} - \boldsymbol{l}_{\perp 1}) L_0(z, \boldsymbol{l}_{\perp 1})\, d\boldsymbol{l}_{\perp 1} \cdots d\boldsymbol{l}_{\perp n}\,. \tag{5.34}$$

Here $\mathbf{l}_\perp$ is the projection of $\hat{\boldsymbol{l}}$ on the plane tangent to the sphere in $\hat{\boldsymbol{l}}_0$.

As the result of the conversion to the Fourier transform of $L_0(z, \hat{\boldsymbol{l}})$ and $p(\mathbf{l}_\perp - \mathbf{l}'_\perp)$ we obtain for plane unidirectional source on the basis of the convolution theorem:

$$L(\tau, \boldsymbol{l}_\perp) = \mathrm{e}^{-\tau/\mu_0} \int_{-\infty}^{+\infty} L_0(k) \exp\left(\frac{\omega_0 \tau p(k)}{\mu_0}\right) J_0(k l_\perp) k\, dk\,, \tag{5.35}$$

where $J_0(\cdot)$ is the zero-order Bessel function.

In Wang and Guth (1951), the relationship between the three SA forms and the comparison of their accuracy were considered. It was shown that the second SA form is the most consistent and accurate, since it neglects only the ray path variance. However, it does not allow generalizing for other sources since, due to the symmetry of plane unidirectional source, $L_0(\hat{\boldsymbol{l}})$ does not depend on ξ. The third SA form could be derived from the second one for small angles when it is possible to replace rotations on a sphere with plane translation:

$$P_k(\cos\vartheta) \approx J_0(k\vartheta), \quad \sum_{k=0}^{\infty} \frac{2k+1}{4\pi} \to \frac{1}{2\pi} \int_{-\infty}^{+\infty} k\, dk\,. \tag{5.36}$$

The transition from the second and third SA forms to the first one is based on the expansion of p_k or $p(k)$ in a Taylor series with the preservation of the first three terms. It should be noted that the presentation (5.26) is not compatible with the exact boundary conditions (Wang ad Guth, 1951), and SA neglects backscattering in the boundary conditions.

It can be seen from this review that the first SA form (Bothe, 1929) cannot be used for removing the anisotropic part of the solution, since it smooths all singularities of the exact solution. The third SA form is the most flexible: in its frameworks, the solution for all elementary sources was achieved and the generalization in the case of polarization was carried out (Muehlschlegel and Koppe, 1958). However, its analytical form as a Hankel transformation results in an extremely inconvenient source function (5.10) that makes finding the smooth part of (5.9) a more complex problem than the original boundary value problem (5.7).

The second SA form (Goudsmit and Saunderson, 1940) has a simple analytical form that does not complicate the original boundary value problem but in the form developed by Goudsmit and Saunderson (1940) it is acceptable only for solution of plane unidirectional source problem. The final form of MSH developed in (5.11)–(5.24) give the same results as those developed by Goudsmit and Saunderson (1940) for the case of the plane unidirectional source. However, the approach itself is easily generalized in the case of the arbitrary geometry with polarization.

We show that MSH contains all the singularities of the exact solution. Thereto we expand the expression for $Z_k(\tau)$ in (5.22) into a Taylor series in terms of powers of ω_0 that corresponds to the radiance expansion in terms of orders of scattering. Then for the first two orders of scattering we have

$$\begin{aligned} L_a^{(0)}(\tau,\mu,\varphi) &= \sum_{l=0}^{\infty}\sum_{n=-l}^{l}\frac{2l+1}{4\pi}Q_l^n(\mu_0)Q_l^n(\mu)\,\mathrm{e}^{in\varphi}\,\mathrm{e}^{-\tau/\mu_0} \\ &= \delta(\mu-\mu_0)\delta(\varphi)\,\mathrm{e}^{-\tau/\mu_0}, \end{aligned} \tag{5.37}$$

$$\begin{aligned} L_a^{(5.1)}(\tau,\mu,\varphi) &= \frac{\omega_0\tau}{\mu_0}\sum_{l=0}^{\infty}\sum_{n=-l}^{l}\frac{2l+1}{4\pi}p_l Q_l^n(\mu_0)Q_l^n(\mu)\,\mathrm{e}^{in\varphi}\,\mathrm{e}^{-\tau/\mu_0} \\ &= \frac{\omega_0\tau}{4\pi\mu_0}p(\hat{\boldsymbol{l}}_0,\hat{\boldsymbol{l}})\,\mathrm{e}^{-\tau/\mu_0}. \end{aligned} \tag{5.38}$$

It is easy to see that (5.37) coincides completely with the direct nonscattered radiation of a plane unidirectional source containing the angular δ-singularity. For the analysis of (5.38) we write the exact expression of the single scattering radiance for the case of a plane unidirectional source (Van de Hulst, 1948):

$$L_1(\tau,\hat{\boldsymbol{l}}) = \begin{cases} \dfrac{\omega_0}{4\pi}\,\mathrm{e}^{-\tau/\mu_0}\dfrac{\mu_0}{\mu_0-\mu}p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0)\left[1-\exp\left(\tau\dfrac{\mu-\mu_0}{\mu\mu_0}\right)\right], & \mu\geq 0; \\ \dfrac{\omega_0}{4\pi}\,\mathrm{e}^{-\tau/\mu_0}\dfrac{\mu_0}{\mu_0-\mu}p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0)\left[\exp\left(\tau\dfrac{\mu-\mu_0}{\mu\mu_0}\right)-\exp\left(\tau_0\dfrac{\mu-\mu_0}{\mu\mu_0}\right)\right], & \mu<0. \end{cases} \tag{5.39}$$

Neglecting the scattered ray path variance that is equivalent in (5.37) to the passage to the limit $\tau\frac{\mu-\mu_0}{\mu\mu_0}\to 0$ we get

$$L_1(\tau,\mu) = \begin{cases} \dfrac{\omega_0\tau}{4\pi\mu}\,\mathrm{e}^{-\tau/\mu_0}p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0), & \mu\geq 0 \\ \dfrac{\omega_0(\tau-\tau_0)}{4\pi\mu}\,\mathrm{e}^{-\tau/\mu_0}p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0), & \mu<0 \end{cases} \tag{5.40}$$

that has the same result as (5.38) in neglecting the scattered ray path variance $\mu \approx \mu_0$.

A similar approach is valid for higher orders of scattering. All of them result from the exact expressions in neglecting the scattered ray path variance. Therefore, it is reasonable to expect that, after subtracting the regular part of the solution, $\tilde{L}(\tau, \mu, \varphi)$ will be a very smoothly varying function of the sighting angle. And the solution of the boundary value problem will not present difficulties for any numerical method: SH, the discrete coordinate method (DOM), Monte Carlo etc. – a review of almost every known method is given in Lenoble (1985) and Sushkevich (2005). To solve the boundary value problem (5.9) we select the DOM that has split-hair accuracy, has a clear ray interpretation that allows a simple generalization in the case of arbitrary boundary conditions and allows us to use analytical methods in the case of a plane unidirectional source. Such an approach, developed on the basis of the Goudsmith–Saunderson approximation and the numerical solution (5.9) by the DOM was for the first time proposed by Romanova (1962a,b) and was widely used by Irvine (1968). Unlike them, we determine an anisotropic part on the basis of MSH that allows us to extend the approach to the arbitrary medium geometry, to include polarization effects, and to get an analytical solution in the case of a plane unidirectional source.

5.5 Determination of the solution of the regular part in a plane unidirectional source problem

To get the regular part of the solution it is necessary to solve the RTE boundary problem (5.9) with the source function $S(\tau, \mu, \varphi)$. It is easy to compute the expression (5.10) for the source function using (5.22), (5.24), and the addition theorem for the spherical functions:

$$S(\tau, \mu, \varphi) = \sum_{m=-\infty}^{\infty} \sum_{k=0}^{K} \frac{2k+1}{4\pi} s_k^m(\tau) Q_k^m(\mu)\, \mathrm{e}^{im\varphi}\,, \tag{5.41}$$

where $d_k = 1 - \omega_0 p_k$,

$$\begin{aligned} s_k^m(\tau) \;=\; & \frac{\sqrt{(k-m+1)(k+m+1)}}{(2k+1)\mu_0}\, d_{k+1} Q_{k+1}^m(\mu_0) Z_{k+1}(\tau) \\ & + \frac{\sqrt{(k-m)(k+m)}}{(2k+1)\mu_0}\, d_{k-1} Q_{k-1}^m Z_{k-1} - d_k Q_k^m Z_k\,. \end{aligned} \tag{5.42}$$

The boundary problem (5.9) for the regular part $\tilde{L}(\tau, \mu, \varphi)$ is solvable by the arbitrary numerical method. However, the most effective method uses the problem plane symmetry as much as possible. Firstly note that the differential transfer operator does not depend on the azimuthal angle φ. Therefore, the unknown function is expanded in Fourier series on azimuth

$$\tilde{L}(\tau, \mu, \varphi) = \sum_{m=-M}^{m=M} C^m(\tau, \mu)\, \mathrm{e}^{im\varphi} = \sum_{m=0}^{m=M} (2 - \delta_{m0}) C^m(\tau, \mu)\, \cos m\varphi. \tag{5.43}$$

Note that azimuthal symmetry is used in (5.43).

With the substitution of (5.43) in the RTE of the boundary problem (5.9), using (5.13) and admitting the orthogonality of the azimuthal harmonics we get the equation:

$$\mu\frac{dC^m(\tau,\mu)}{d\tau} = -C^m(\tau,\mu) + \frac{\omega_0}{2}\sum_{k=m}^{N_x}(2k+1)p_kQ_k^m(\mu)\int\limits_{-1}^{1}Q_k^m(\mu')C^m(\tau,\mu')\,d\mu' + \sum_{k=0}^{\infty}\frac{2k+1}{4\pi}s_k^m(\tau)Q_k^m(\mu)\,. \tag{5.44}$$

Any solution method of the integro-differential equation (5.44) is based on the replacement of an integral with the finite sum. The method of discrete ordinates is the most effective one where an integral is presented as the Gaussian quadrature (Krylov, 2006). At first, it was used in Wick (1943) and then developed by Chandrasekhar (1950) and Thomas and Stamnes (2002). Since the solution regular part $\tilde{L}(\tau,\mu,\varphi)$ is the smooth angular function, it can be approximately presented by the finite set of its N values or ordinates. Note that in case of the homogeneous slab, the radiative transfer is symmetrical with respect to the upward ($\mu<0$) and downward ($\mu>0$) streams of radiation (Stamnes and Swanson, 1981). Therefore, for the best integral representation in the equation (5.44) the Gaussian quadrature in the following form (Stamnes and Swanson, 1981; Sykes, 1951) is used:

$$\int\limits_{-1}^{1}Q_k^m(\mu')C^m(\tau,\mu')\,d\mu' \approx \frac{1}{2}\sum_{j=1}^{N/2}w_jC_j^-(\tau)Q_k^m(\mu_j^-) + \frac{1}{2}\sum_{j=1}^{N/2}w_jC_j^+(\tau)Q_k^m(\mu_j^+)\,, \tag{5.45}$$

where $\mu_j^- = (\mu_j-1)/2$, $\mu_j^+ = (\mu_j+1)/2$, μ_j are zeros of the Legendre polynomials $P_K(\mu)$, w_j are the weighting coefficients of the Gaussian quadrature; $C_j^{\pm}(\tau) \equiv C_j^m(\tau,\mu_j^{\pm})$. Notice that N must be even. The additional advantage of this scheme is the search of only $N/2$ roots.

In this case the set (5.44) could be replaced by the set of N ordinary differential equations

$$\mu_i^{\pm}\frac{dC_i^{\pm}(\tau)}{d\tau} = -C_i^{\pm}(\tau) + \frac{\omega_0}{4}\sum_{j=1}^{N/2}w_j\sum_{k=m}^{N_x}(2k+1)p_kQ_k^m(\mu_i^{\pm})\left(C_j^-(\tau)Q_k^m(\mu_j^-) + C_j^+(\tau)Q_k^m(\mu_j^+)\right) + S_i^{\pm}(\tau), \tag{5.46}$$

where

$$S_i^{\pm}(\tau) = \sum_{k=m}^{K}\frac{2k+1}{4\pi}s_k^m(\tau)Q_k^m(\mu_i)\,,$$

K is the number of terms in the anisotropic part solution presentation. It is difficult to point out the unique selection rule of the value K: it depends both on the scattering anisotropy and on the absorption in the medium determined by ω_0. However, it is always $K > N_x$.

Let values of $\{\mu_i^{\pm}\}$ be in ascending order of the cosine modulus of the sighting angle. We define the column vectors

$$\vec{C}_+ = \begin{bmatrix} C_1^+ \\ \vdots \\ C_{N/2}^+ \end{bmatrix}, \quad \vec{C}_- = \begin{bmatrix} C_1^- \\ \vdots \\ C_{N/2}^- \end{bmatrix},$$
$$\vec{S}_+ = \begin{bmatrix} S_1^+ / \mu_1^+ \\ \vdots \\ S_{N/2}^+ \Big/ \mu_{N/2}^+ \end{bmatrix}, \quad \vec{S}_- = \begin{bmatrix} S_1^- / \mu_1^- \\ \vdots \\ S_{N/2}^- \Big/ \mu_{N/2}^- \end{bmatrix}, \tag{5.47}$$

and the matrices

$$\overleftrightarrow{X} = \frac{\omega_0}{4} \left[\sum_{k=m}^{K} (2k+1) p_k Q_k^m(\mu_i^{\pm}) Q_k^m(\mu_j^+) \quad \sum_{k=m}^{K} (2k+1) p_k Q_k^m(\mu_i^{\pm}) Q_k^m(\mu_j^-) \right],$$
$$\overleftrightarrow{M} = \begin{bmatrix} \mu_i^+ & 0 \\ 0 & \mu_i^- \end{bmatrix}, \quad \overleftrightarrow{W} = \mathrm{diag}(w_i)\,,$$

that allows rewriting (5.46) in the matrix form

$$\frac{d}{d\tau} \begin{bmatrix} \vec{C}_+ \\ \vec{C}_- \end{bmatrix} = -\overleftrightarrow{M}^{-1} (\overleftrightarrow{1} - \overleftrightarrow{X}\overleftrightarrow{W}) \begin{bmatrix} \vec{C}_+ \\ \vec{C}_- \end{bmatrix} + \begin{bmatrix} \vec{F}_+ \\ \vec{F}_- \end{bmatrix}. \tag{5.48}$$

From here on the over symbol '$\rightarrow$' denotes a column vector and the over symbol '$\leftrightarrow$' denotes a square matrix of the corresponding size.

Now the matrix has the following structure:

$$\overleftrightarrow{X} = \begin{bmatrix} x_1 & x_2 \\ x_2 & x_1 \end{bmatrix}, \quad \overleftrightarrow{W} = \begin{bmatrix} w & 0 \\ 0 & w \end{bmatrix}, \quad \overleftrightarrow{M}^{-1} = \begin{bmatrix} m & 0 \\ 0 & -m \end{bmatrix}, \quad m = \left\{ \frac{1}{\mu_i^+} \right\}.$$

This allows us to write down expressions

$$\overleftrightarrow{X}\overleftrightarrow{W} = \begin{bmatrix} x_1 w & x_2 w \\ x_2 w & x_1 w \end{bmatrix},$$
$$\overleftrightarrow{M}^{-1} (\overleftrightarrow{1} - \overleftrightarrow{X}\overleftrightarrow{W}) = \begin{bmatrix} m(1 - x_1 w) & -m x_2 w \\ m x_2 w & -m(1 - x_1 w) \end{bmatrix} = \begin{bmatrix} \overleftrightarrow{\alpha} & -\overleftrightarrow{\beta} \\ \overleftrightarrow{\beta} & -\overleftrightarrow{\alpha} \end{bmatrix} \equiv \overleftrightarrow{B}\,. \tag{5.49}$$

From (5.49) and (5.46) the matrix equation (5.48) can be rewritten as:

$$\begin{cases} \dfrac{d\vec{C}_+}{d\tau} = -\alpha \vec{C}_+ + \beta \vec{C}_+ + \vec{S}_+\,, \\ \dfrac{d\vec{C}_-}{d\tau} = -\beta \vec{C}_+ + \alpha \vec{C}_- + \vec{S}_-\,. \end{cases} \tag{5.50}$$

The solution of the set (5.50) has the form

$$-\begin{bmatrix}\vec{C}_+(0)\\ \vec{C}_-(0)\end{bmatrix}+\mathrm{e}^{\overleftrightarrow{B}\tau_0}\begin{bmatrix}\vec{C}_+(\tau_0)\\ \vec{C}_-(\tau_0)\end{bmatrix}=\int\limits_0^{\tau_0}\mathrm{e}^{\overleftrightarrow{B}t}\begin{bmatrix}\overleftrightarrow{S}_+(t)\\ \overleftrightarrow{S}_-(t)\end{bmatrix}dt\,. \tag{5.51}$$

The set (5.51) contains N equations with $2N$ unknowns. Missing N equations can be found from the boundary conditions:

$$C(0,\mu_i)|_{\forall i:\ \mu_i\geq 0}=0,\quad C(\tau_0,\mu_i)|_{\forall i:\ \mu_i\leq 0}=-L_a(\tau_0,\mu_i)\,, \tag{5.52}$$

or in the matrix form

$$\begin{bmatrix}\overleftrightarrow{1} & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\vec{C}_+(0)\\ \vec{C}_-(0)\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix},\quad \begin{bmatrix}0 & 0\\ 0 & \overleftrightarrow{1}\end{bmatrix}\begin{bmatrix}\vec{C}_+(\tau_0)\\ \vec{C}_-(\tau_0)\end{bmatrix}=-\begin{bmatrix}0\\ \vec{L}_a(\tau_0)\end{bmatrix},$$
$$\vec{L}_a(\tau_0)\equiv\{L_a(\tau_0,\mu_i)\}\ . \tag{5.53}$$

The item $L_a(\tau_0,\mu_i)$ at the bottom boundary condition is connected with the physical lack of fit of MSH solutions because of the neglect of the backscattering.

The matrix exponent could be represented in the form (Sykes, 1951)

$$\mathrm{e}^{\overleftrightarrow{B}\tau}=\overleftrightarrow{U}\mathrm{e}^{\overleftrightarrow{\Gamma}\tau}\overleftrightarrow{U}^{-1}, \tag{5.54}$$

where

$$\overleftrightarrow{\Gamma}=\begin{bmatrix}\overleftrightarrow{\Gamma}_- & 0\\ 0 & \overleftrightarrow{\Gamma}_+\end{bmatrix},\ \overleftrightarrow{\Gamma}_\pm=\mathrm{diag}\{\gamma_i\},\ \text{and}\ \mathrm{e}^{\overleftrightarrow{\Gamma}\tau}=\begin{bmatrix}\mathrm{e}^{\overleftrightarrow{\Gamma}_-\tau} & 0\\ 0 & \mathrm{e}^{\overleftrightarrow{\Gamma}_+\tau}\end{bmatrix},\ \{\gamma_i\}$$

are the eigenvalues that are sorted by ascending $\gamma_i<\gamma_{i+1}$, $\overleftrightarrow{U}$ is the eigenvector matrix of the matrix $\overleftrightarrow{B}$.

Taking into account (5.49), for the matrix $\overleftrightarrow{B}$ it is possible to state the eigenvalue problem (Stamnes and Swanson, 1981)

$$\begin{bmatrix}\overleftrightarrow{\alpha} & -\overleftrightarrow{\beta}\\ \overleftrightarrow{\beta} & -\overleftrightarrow{\alpha}\end{bmatrix}\begin{bmatrix}\vec{u}_+\\ \vec{u}_-\end{bmatrix}=\gamma\begin{bmatrix}\vec{u}_+\\ \vec{u}_-\end{bmatrix}, \tag{5.55}$$

that allows rewriting it differently as the equation set

$$\begin{cases}\overleftrightarrow{\alpha}\vec{u}_+-\overleftrightarrow{\beta}\vec{u}_-=\gamma\vec{u}_+,\\ \overleftrightarrow{\beta}\vec{u}_+-\overleftrightarrow{\alpha}\vec{u}_-=\gamma\vec{u}_-.\end{cases} \tag{5.56}$$

Let us add and subtract in pairs the equations of the set (5.56) that results in the new equation set

$$\begin{cases}(\overleftrightarrow{\alpha}+\overleftrightarrow{\beta})(\vec{u}_+-\vec{u}_-)=\gamma(\vec{u}_++\vec{u}_-)\,,\\ (\overleftrightarrow{\alpha}-\overleftrightarrow{\beta})(\vec{u}_++\vec{u}_-)=\gamma(\vec{u}_+-\vec{u}_-)\,.\end{cases} \tag{5.57}$$

By expressing $(\vec{u}_+ - \vec{u}_-)$ from the lower equation and by substituting it in the upper one we get the eigenvalues problem

$$(\overleftrightarrow{\alpha} + \overleftrightarrow{\beta})(\overleftrightarrow{\alpha} - \overleftrightarrow{\beta})(\vec{u}_+ + \vec{u}_-) = \gamma^2(\vec{u}_+ + \vec{u}_-)\,. \tag{5.58}$$

We define the eigenvalues γ and corresponding eigenvectors $(\vec{u}_+ + \vec{u}_-)$ from equation (5.58). Using the obtained values from the lower equation of the set (5.57) we define the eigenvectors $(\vec{u}_+ - \vec{u}_-)$ that allow us to compute the matrices $\overleftrightarrow{\Gamma}$ and $\overleftrightarrow{U}$.

Now, using (5.40) we rewrite the set (5.51) as:

$$-\begin{bmatrix} \vec{C}_+(0) \\ \vec{C}_-(0) \end{bmatrix} + \overleftrightarrow{U}\,\mathrm{e}^{\overleftrightarrow{\Gamma}\tau_0}\overleftrightarrow{U}^{-1}\begin{bmatrix} \vec{C}_+(\tau_0) \\ \vec{C}_-(\tau_0) \end{bmatrix} = \overleftrightarrow{U}\sum_{k=0}^{\infty}\frac{2k+1}{4\pi}\int_0^{\tau_0}\mathrm{e}^{\overleftrightarrow{\Gamma}t}s_k^m(t)\,dt\overleftrightarrow{U}^{-1}\begin{bmatrix} \vec{q}_+ \\ \vec{q}_- \end{bmatrix}, \tag{5.59}$$

where

$$\begin{bmatrix} \vec{q}_+ \\ \vec{q}_- \end{bmatrix} = \left[Q_k^m(\mu_i^{\pm})\right]\,.$$

If we introduce the designation

$$\overleftrightarrow{i}_k \equiv \int_0^{\tau_0}\exp\left[-\left(\frac{d_k}{\mu_0} - \overleftrightarrow{\Gamma}\right)t\right]\,dt = \left[\overleftrightarrow{1} - \exp\left[-\left(\frac{d_k}{\mu_0}\overleftrightarrow{1} - \overleftrightarrow{\Gamma}\right)\tau_0\right]\right]\left(\frac{d_k}{\mu_0}\overleftrightarrow{1} - \overleftrightarrow{\Gamma}\right)^{-1}, \tag{5.60}$$

the integral in equation (5.59) can be represented by the expression

$$\begin{aligned}\int_0^{\tau_0}\mathrm{e}^{\overleftrightarrow{\Gamma}t}s_k^m(t)\,dt \;=\; & \frac{1}{(2k+1)\mu_0}\left[\sqrt{(k-m+1)(k+m+1)}d_{k+1}Q_{k+1}^m(\mu_0)\overleftrightarrow{i}_{k+1}(\tau)\right. \\ & \left.+\sqrt{(k-m)(k+m)}d_{k-1}Q_{k-1}^m(\mu_0)\overleftrightarrow{i}_{k-1}\right] \\ & - d_kQ_k^m(\mu_0)\overleftrightarrow{i}_k \equiv \overleftrightarrow{I}_k(\tau_0).\end{aligned} \tag{5.61}$$

The system matrix conditionality decreases quickly with the medium depth. To eliminate this effect we use the scale transformation (Karp et al., 1980) and multiply (5.59) by matrix $\overleftrightarrow{S}\overleftrightarrow{U}^{-1}$ that reduces the system to the form

$$-\overleftrightarrow{S}\overleftrightarrow{U}^{-1}\begin{bmatrix} \vec{C}_+(0) \\ \vec{C}_-(0) \end{bmatrix} + \overleftrightarrow{H}\overleftrightarrow{U}^{-1}\begin{bmatrix} \vec{C}_+(\tau_0) \\ \vec{C}_-(\tau_0) \end{bmatrix} = \sum_{k=0}^{\infty}\frac{2k+1}{4\pi}\overleftrightarrow{S}\overleftrightarrow{I}_k(\tau_0)\overleftrightarrow{U}^{-1}\begin{bmatrix} \vec{q}_+ \\ \vec{q}_- \end{bmatrix}, \tag{5.62}$$

where

$$\overleftrightarrow{S} = \begin{bmatrix} 0 & \mathrm{e}^{-\overleftrightarrow{\Gamma}_+\tau_0} \\ \overleftrightarrow{1} & 0 \end{bmatrix}, \quad \overleftrightarrow{H} = \overleftrightarrow{S}\mathrm{e}^{\overleftrightarrow{\Gamma}\tau_0} = \begin{bmatrix} 0 & \mathrm{e}^{-\overleftrightarrow{\Gamma}_+\tau_0} \\ \overleftrightarrow{1} & 0 \end{bmatrix}\begin{bmatrix} \mathrm{e}^{\overleftrightarrow{\Gamma}_-\tau_0} & 0 \\ 0 & \mathrm{e}^{\overleftrightarrow{\Gamma}_+\tau_0} \end{bmatrix} = \begin{bmatrix} 0 & \overleftrightarrow{1} \\ \mathrm{e}^{\overleftrightarrow{\Gamma}_-\tau_0} & 0 \end{bmatrix}, \tag{5.63}$$

and the integrals in the multiplier $\overleftrightarrow{S}\overleftrightarrow{I}_k(\tau_0)$ are:

$$\overleftrightarrow{S}\overleftrightarrow{i}_k(\tau_0) = \begin{bmatrix} \overleftrightarrow{0} & \left(\mathrm{e}^{-\overleftrightarrow{\Gamma}_+\tau_0} - \overleftrightarrow{1}\mathrm{e}^{-d_k\tau_0/\mu_0}\right)\left(\frac{d_k}{\mu_0}\overleftrightarrow{1} - \overleftrightarrow{\Gamma}_+\right)^{-1} \\ \left(\overleftrightarrow{1} - \mathrm{e}^{\overleftrightarrow{\Gamma}_-\tau_0}\mathrm{e}^{-d_k\tau_0/\mu_0}\right)\left(\frac{d_k}{\mu_0}\overleftrightarrow{1} - \overleftrightarrow{\Gamma}_-\right)^{-1} & \overleftrightarrow{0} \end{bmatrix}. \tag{5.64}$$

The equation set (5.62) with the boundary conditions (5.53) can be reduced and written only with respect to reflected $\vec{C}_-(0)$ and transmitted $\vec{C}_+(\tau_0)$ components. Let us open the left-hand members in the equation set

$$\begin{aligned} -\overleftrightarrow{S}\overleftrightarrow{U}^{-1}\vec{C}(0) + \overleftrightarrow{H}\overleftrightarrow{U}^{-1}\vec{C}(\tau_0) \;=\; & -\begin{bmatrix} \mathrm{e}^{-\overleftrightarrow{\Gamma}_+\tau_0}\overleftrightarrow{u}_{21} & \mathrm{e}^{-\overleftrightarrow{\Gamma}_+\tau_0}\overleftrightarrow{u}_{22} \\ \overleftrightarrow{u}_{11} & \overleftrightarrow{u}_{12} \end{bmatrix}\begin{bmatrix} \vec{C}_+(0) \\ \vec{C}_-(0) \end{bmatrix} \\ & + \begin{bmatrix} \overleftrightarrow{u}_{21} & \overleftrightarrow{u}_{22} \\ \mathrm{e}^{\overleftrightarrow{\Gamma}_-\tau_0}\overleftrightarrow{u}_{11} & \mathrm{e}^{\overleftrightarrow{\Gamma}_-\tau_0}\overleftrightarrow{u}_{12} \end{bmatrix}\begin{bmatrix} \vec{C}_+(\tau_0) \\ \vec{C}_-(\tau_0) \end{bmatrix}, \end{aligned} \tag{5.65}$$

that changes set (5.62) to

$$\begin{aligned} & \begin{bmatrix} -\mathrm{e}^{-\overleftrightarrow{\Gamma}_+\tau_0}\overleftrightarrow{u}_{22} & \overleftrightarrow{u}_{21} \\ -\overleftrightarrow{u}_{12} & \mathrm{e}^{\overleftrightarrow{\Gamma}_-\tau_0}\overleftrightarrow{u}_{11} \end{bmatrix}\begin{bmatrix} \vec{C}_-(0) \\ \vec{C}_+(\tau_0) \end{bmatrix} \\ & \quad = \begin{bmatrix} \vec{J}_+ \\ \vec{J}_- \end{bmatrix} + \begin{bmatrix} \mathrm{e}^{-\overleftrightarrow{\Gamma}_+\tau_0}\overleftrightarrow{u}_{21} & -\overleftrightarrow{u}_{22} \\ \overleftrightarrow{u}_{11} & -\mathrm{e}^{\overleftrightarrow{\Gamma}_-\tau_0}\overleftrightarrow{u}_{12} \end{bmatrix}\begin{bmatrix} \vec{C}_+(0) \\ \vec{C}_-(\tau_0) \end{bmatrix}, \end{aligned} \tag{5.66}$$

where

$$\overleftrightarrow{U}^{-1} \equiv \begin{bmatrix} \overleftrightarrow{u}_{11} & \overleftrightarrow{u}_{12} \\ \overleftrightarrow{u}_{21} & \overleftrightarrow{u}_{22} \end{bmatrix}, \quad \begin{bmatrix} \vec{J}_+ \\ \vec{J}_- \end{bmatrix} \equiv \sum_{k=0}^{\infty}\frac{2k+1}{4\pi}\overleftrightarrow{S}\overleftrightarrow{I}_k(\tau_0)\overleftrightarrow{U}^{-1}\begin{bmatrix} \vec{q}_+ \\ \vec{q}_- \end{bmatrix}.$$

The equation set (5.66) can be simplified at the analysis of the practically important case of the reflection from the semi-infinite slab – 'ocean'. In this case $\tau_0 \to \infty$ and the relations are

$$\mathrm{e}^{-\overleftrightarrow{\Gamma}_+\tau_0} \to 0, \;\; \mathrm{e}^{\overleftrightarrow{\Gamma}_-\tau_0} \to 0, \;\; \vec{C}_+(\tau_0) \to 0, \;\; \vec{C}_-(\tau_0) = -\vec{L}_a(\tau_0) \to 0, \;\; \overleftrightarrow{J}_-(\tau_0) \to 0\,. \tag{5.67}$$

So the solution gets a very simple form

$$\vec{C}_-(0) = -\overleftrightarrow{u}_{12}^{-1}\vec{J}_- = \frac{-\overleftrightarrow{u}_{12}^{-1}}{4\pi\mu_0}\sum_{k=0}^{N_z}(1-\omega_0 p_k)Q_k^m(\mu_0)\overleftrightarrow{I}_-\left(\overleftrightarrow{u}_{11}\vec{m}_-\vec{q}_- + \overleftrightarrow{u}_{12}\vec{m}_+\vec{q}_+\right), \tag{5.68}$$

where

$$\begin{aligned} \overleftrightarrow{I}_- \;=\; & \sqrt{(k-m+1)(k+m+1)}Q_{k+1}^m(\mu_0)\overleftrightarrow{j}_{k+1} \\ & + \sqrt{(k-m)(k+m)}Q_{k-1}^m(\mu_0)\overleftrightarrow{j}_{k-1} - (2k+1)Q_k^m(\mu_0)\overleftrightarrow{j}_k, \;\; \overleftrightarrow{j}_k \\ =\; & d_k\left(d_k\overleftrightarrow{1} - \mu_0\overleftrightarrow{\Gamma}_-\right)^{-1}, \quad \vec{m}_\pm = \mathrm{diag}\left(1/\mu_i^\pm\right). \end{aligned}$$

It is obvious that the expression (5.68) is the analytical solution of the Ambartsumian problem of radiation reflection from a semi-infinite medium at an arbitrary phase function for the discrete RTE (5.46).

Solving equation (5.66) with respect to reflected $\vec{C}_-(0)$ and transmitted $\vec{C}_+(\tau_0)$ components we get

$$\begin{bmatrix} \vec{C}_-(0) \\ \vec{C}_+(\tau_0) \end{bmatrix} = \begin{bmatrix} \vec{F}_- \\ \vec{F}_+ \end{bmatrix} + \begin{bmatrix} \overleftrightarrow{R}_- & \overleftrightarrow{T}_- \\ \overleftrightarrow{T}_+ & \overleftrightarrow{R}_+ \end{bmatrix} \begin{bmatrix} \vec{C}_+(0) \\ \vec{C}_-(\tau_0) \end{bmatrix}, \tag{5.69}$$

where

$$\begin{bmatrix} \vec{F}_- \\ \vec{F}_+ \end{bmatrix} \equiv \overleftrightarrow{A}^{-1} \begin{bmatrix} \vec{J}_+ \\ \vec{J}_- \end{bmatrix}, \quad \begin{bmatrix} \overleftrightarrow{R}_- & \overleftrightarrow{T}_- \\ \overleftrightarrow{T}_+ & \overleftrightarrow{R}_+ \end{bmatrix} \equiv \overleftrightarrow{A}^{-1} \begin{bmatrix} \mathrm{e}^{-\overleftrightarrow{\Gamma}_+ \tau_0} \overleftrightarrow{u}_{21} & -\overleftrightarrow{u}_{22} \\ \overleftrightarrow{u}_{11} & -\mathrm{e}^{\overleftrightarrow{\Gamma}_- \tau_0} \overleftrightarrow{u}_{12} \end{bmatrix},$$

$$\overleftrightarrow{A}^{-1} \equiv \begin{bmatrix} -\mathrm{e}^{-\overleftrightarrow{\Gamma}_+ \tau_0} \overleftrightarrow{u}_{22} & \overleftrightarrow{u}_{21} \\ -\overleftrightarrow{u}_{12} & \mathrm{e}^{\overleftrightarrow{\Gamma}_- \tau_0} \overleftrightarrow{u}_{11} \end{bmatrix}.$$

The expression (5.69) defines the radiative transfer through the slab: the reflected $\vec{C}_-(0)$ and the transmitted $\vec{C}_+(\tau_0)$ radiation are expressed by the radiation incident on the slab border from above $\vec{C}_+(0)$ and from below $\vec{C}_-(\tau_0)$. The merit of (5.69) is the possibility of translation to a vertically inhomogeneous, stratified slab dividing it into an arbitrary quantity of homogeneous slabs. In this case, two adjoining slabs can be replaced by one slab that is described by the equivalent expression (5.69) that has effective parameters expressed by the original slab parameters. Thereby, the property of the radiative transfer by a slab possesses invariance (Stokes, 1862), and the matrix elements of the united slab are determined with the help of the matrix-operator method (Plass et al., 1973) .

We consider the case of the two-layer medium from two adjacent slabs

$$\begin{bmatrix} \vec{C}^1_- \\ \vec{C}^1_+ \end{bmatrix} = \begin{bmatrix} \vec{F}^1_- \\ \vec{F}^1_+ \end{bmatrix} + \begin{bmatrix} \overleftrightarrow{R}_{1-} & \overleftrightarrow{T}_{1-} \\ \overleftrightarrow{T}_{1+} & \overleftrightarrow{R}_{1+} \end{bmatrix} \begin{bmatrix} \vec{C}^1_\downarrow \\ \vec{C}^1_\uparrow \end{bmatrix},$$

$$\begin{bmatrix} \vec{C}^2_- \\ \vec{C}^2_+ \end{bmatrix} = \begin{bmatrix} \vec{F}^2_- \\ \vec{F}^2_+ \end{bmatrix} + \begin{bmatrix} \overleftrightarrow{R}_{2-} & \overleftrightarrow{T}_{2-} \\ \overleftrightarrow{T}_{2+} & \overleftrightarrow{R}_{2+} \end{bmatrix} \begin{bmatrix} \vec{C}^2_\downarrow \\ \vec{C}^2_\uparrow \end{bmatrix}, \tag{5.70}$$

where the subscript defines an upper (5.1) or lower (5.2) slab. The radiation incident on the slab from above or from below is designated by the vertical arrows.

Note that since the slabs are adjoining

$$\vec{C}^1_+ = \vec{C}^2_\downarrow \equiv \vec{C}_\downarrow, \quad \vec{C}^2_- = \vec{C}^1_\uparrow \equiv \vec{C}_\uparrow, \tag{5.71}$$

we could rewrite (5.70) as

$$\begin{bmatrix} \vec{C}^1_- \\ \vec{C}_\downarrow \end{bmatrix} = \begin{bmatrix} \vec{F}^1_- \\ \vec{F}^1_+ \end{bmatrix} + \begin{bmatrix} \overleftrightarrow{R}_{1-} & \overleftrightarrow{T}_{1-} \\ \overleftrightarrow{T}_{1+} & \overleftrightarrow{R}_{1+} \end{bmatrix} \begin{bmatrix} \vec{C}^1_\downarrow \\ \vec{C}_\uparrow \end{bmatrix},$$

$$\begin{bmatrix} \vec{C}_\uparrow \\ \vec{C}^2_+ \end{bmatrix} = \begin{bmatrix} \vec{F}^2_- \\ \vec{F}^2_+ \end{bmatrix} + \begin{bmatrix} \overleftrightarrow{R}_{2-} & \overleftrightarrow{T}_{2-} \\ \overleftrightarrow{T}_{2+} & \overleftrightarrow{R}_{2+} \end{bmatrix} \begin{bmatrix} \vec{C}_\downarrow \\ \vec{C}^2_\uparrow \end{bmatrix}. \tag{5.72}$$

Solving equation set (5.70) for reflected $\vec{C}^1_-$ and transmitted $\vec{C}^2_+$ radiation through the radiation incident on the slab $\vec{C}^1_\downarrow$ and $\vec{C}^2_\uparrow$ we will get:

$$\vec{C}^1_- = \vec{F}^1_- + \overleftrightarrow{T}_{1-}\left(\overleftrightarrow{1} - \overleftrightarrow{R}_{2-}\overleftrightarrow{R}_{1+}\right)^{-1}\overleftrightarrow{R}_{2-}\vec{F}^1_+ + \overleftrightarrow{T}_{1-}\left(\overleftrightarrow{1} - \overleftrightarrow{R}_{2-}\overleftrightarrow{R}_{1+}\right)^{-1}\vec{F}^2_- + \left[\overleftrightarrow{R}_{1-} + \overleftrightarrow{T}_{1-}\left(\overleftrightarrow{1} - \overleftrightarrow{R}_{2-}\overleftrightarrow{R}_{1+}\right)^{-1}\overleftrightarrow{R}_{2-}\overleftrightarrow{T}_{1+}\right]\vec{C}^1_\downarrow + \overleftrightarrow{T}_{1-}\left(\overleftrightarrow{1} - \overleftrightarrow{R}_{2-}\overleftrightarrow{R}_{1+}\right)^{-1}\overleftrightarrow{T}_{2-}\vec{C}^2_\uparrow \tag{5.73}$$

$$\vec{C}^2_+ = \overleftrightarrow{T}_{2+}\left(\overleftrightarrow{1} - \overleftrightarrow{R}_{1+}\overleftrightarrow{R}_{2-}\right)^{-1}\vec{F}^1_+ + \overleftrightarrow{T}_{2+}\left(\overleftrightarrow{1} - \overleftrightarrow{R}_{1+}\overleftrightarrow{R}_{2-}\right)^{-1}\overleftrightarrow{R}_{1+}\vec{F}^2_- + \vec{F}^2_+ + \overleftrightarrow{T}_{2+}\left(\overleftrightarrow{1} - \overleftrightarrow{R}_{1+}\overleftrightarrow{R}_{2-}\right)^{-1}\overleftrightarrow{T}_{1+}\vec{C}^1_\downarrow + \left[\overleftrightarrow{R}_{2+} + \overleftrightarrow{T}_{2+}\left(\overleftrightarrow{1} - \overleftrightarrow{R}_{1+}\overleftrightarrow{R}_{2-}\right)^{-1}\overleftrightarrow{R}_{1+}\overleftrightarrow{T}_{2-}\right]\vec{C}^2_\uparrow \tag{5.74}$$

By introducing designations

$$\overleftrightarrow{\alpha} = \left(\overleftrightarrow{1} - \overleftrightarrow{R}_{2-}\overleftrightarrow{R}_{1+}\right)^{-1}, \tag{5.75}$$

we write (5.73)–(5.74) in the matrix form

$$\begin{bmatrix}\vec{C}^1_- \\ \vec{C}^2_+\end{bmatrix} = \begin{bmatrix}\vec{F}^1_- + \overleftrightarrow{T}_{1-}\overleftrightarrow{\alpha}\left(\overleftrightarrow{R}_{2-}\vec{F}^1_+ + \vec{F}^2_-\right) \\ \overleftrightarrow{T}_{2+}\overleftrightarrow{\alpha}\left(\vec{F}^1_+ + \overleftrightarrow{R}_{1+}\vec{F}^2_-\right) + \vec{F}^2_+\end{bmatrix} + \begin{bmatrix}\overleftrightarrow{R}_{1-} + \overleftrightarrow{T}_{1-}\overleftrightarrow{\alpha}\overleftrightarrow{R}_{2-}\overleftrightarrow{T}_{1+} & \overleftrightarrow{T}_{1-}\overleftrightarrow{\alpha}\overleftrightarrow{T}_{2-} \\ \overleftrightarrow{T}_{2+}\overleftrightarrow{\alpha}\overleftrightarrow{T}_{1+} & \overleftrightarrow{R}_{2+} + \overleftrightarrow{T}_{2+}\overleftrightarrow{\alpha}\overleftrightarrow{R}_{1+}\overleftrightarrow{T}_{2-}\end{bmatrix}\begin{bmatrix}\vec{C}^1_\downarrow \\ \vec{C}^2_\uparrow\end{bmatrix}. \tag{5.76}$$

To count the multilayer systems one has to replace the expression for the angular spectrum of the anisotropic part in (5.20) by the expression

$$Z_k(\tau) = Z_{0k}\exp\left\{-\frac{(1-\omega_0 p_k)\tau}{\mu_0}\right\}, \tag{5.77}$$

where Z_{0k} are the expansion coefficients on the Legendre polynomials of RAD on the upper slab border. In the case of a plane unidirectional source on the upper border, $Z_{0k} = 1$. In case of the multilayer slab, the lower layer in MSH is irradiated from above:

$$Z_{0ik} = \exp\left\{-\frac{(1-\omega_0^{(i-1)}p_k^{(i-1)})\tau_{i-1}}{\mu_0}\right\},$$

where i is the number of the current layer.

5.6 Reflection and transmittance on the boundary of two slabs

At first we handle the special case for the slab illuminated from above by a plane unidirectional source and reflected diffusely from the bottom with the reflectance ρ and get

$$L(\tau_0, \hat{\boldsymbol{l}})\Big|_{\mu<0} = \frac{\rho}{\pi}\int\limits_{(\mu>0)} L(\tau_0, \hat{\boldsymbol{l}}')\mu\, d\hat{\boldsymbol{l}}'. \tag{5.78}$$

If we substitute the radiance representation as a Fourier series (5.43) in the form of a discrete ordinate while the integral (5.78) is presented by a Gaussian quadrature we get

$$C^0(\tau_0,\mu_i^-) = 2\rho\sum_{j=1}^{N/2} C^0(\tau_0,\mu_j^+)\mu_j^+ w_j, \quad \forall m>0:\ C^m(\tau_0,\mu_i^-) = \mathbf{0}\,, \tag{5.79}$$

or in the matrix form

$$m=0:\ \ \vec{C}_-(\tau_0) = 2\rho\overleftrightarrow{R}_L\vec{C}_+(\tau_0); \qquad \forall m>0:\ \ \vec{C}_-(\tau_0) = \mathbf{0}, \tag{5.80}$$

where the Lambertian reflection matrix $\overleftrightarrow{R}_L$ consist of the N/2 same rows $\left\{\mu_j^+ w_j\right\}$.

In correspondence with matrix-operator method (5.76), we get for the reflected component of the zero harmonic:

$$\vec{C}_-(0) = \vec{F}_- + 2\rho\overleftrightarrow{T}_-\left(\overleftrightarrow{1} - 2\rho\overleftrightarrow{R}_L\overleftrightarrow{R}\right)^{-1}\overleftrightarrow{R}_L\vec{F}_+, \tag{5.81}$$

where the superscript 1 relating to the slab is omitted everywhere. All other azimuthal harmonics for $m>0$ are determined by the expression for the single-layer medium (5.69).

However, this approach cannot be straightly applied to the refraction boundary since the ordinate directions ('rays') are changed on the border by the Snell law

$$n_1\,\sin\theta_1 = n_2\,\sin\theta_2\,, \tag{5.82}$$

where n_1, n_2 are refractive indices of media, and the correspondence of ordinate directions is broken. The solution of this problem is given in Nakajima and Tanaka (1986).

Here we handle in detail the refraction for the practically important case, when $n_1 < n_2$. Let us call the first slab 'atmosphere' (every symbol referring to it will be designated by index a) and assume $n_a = 1$; and call the second slab 'ocean' (index o) and assume $n_o > 1$. The cosines of the angles between the rays and the axis OZ in both media in correspondence with (5.82) are related by the expression

$$\mu_a = \sqrt{1 - n_o^2(1-\mu_o^2)}\,. \tag{5.83}$$

It can be seen from (5.83) that for $\mu_o < \mu_t \equiv \sqrt{1-1/n_o^2}$ the total reflection region appears in the ocean when the light rays do not quit the ocean but are perfectly reflected back to the ocean. There are no problems in formulating the boundary conditions in the total reflection region. Let us analyze the integral (5.45) by region:

$$\begin{aligned}\int_{-1}^{1} Q_k^m(\mu')C^m(\tau,\mu')\,d\mu' \ &= \ \int_{-1}^{-\mu_t} Q_k^m(\mu')C^m(\tau,\mu')\,d\mu' + \int_{-\mu_t}^{\mu_t} Q_k^m(\mu')C^m(\tau,\mu')\,d\mu' \\ &\quad + \int_{\mu_t}^{1} Q_k^m(\mu')C^m(\tau,\mu')\,d\mu'\,.\end{aligned} \tag{5.84}$$

The first and the third integrals are connected with the refraction region, and the second is related to the total reflection region. For the second integral, we can make the transformation

$$\int\limits_{-\mu_t}^{\mu_t} Q_k^m(\mu')C^m(\tau,\mu')\,d\mu' = \int\limits_{-1}^{1} Q_k^m(\nu)C^m(\tau,\nu)d\nu, \quad \mu' = \mu_t\nu, \tag{5.85}$$

that makes it possible to use the double Gaussian quadrature with N_t knots and later come over to the two streams of ordinates in this region $\vec{C}_+^t$, $\vec{C}_-^t$, that are interconnected to each other on the border by ideal reflection.

For the first and the last integrals in (5.84), we make the transformation of the integration variable to μ_a by the expression (5.83)

$$\mu_o = \sqrt{1 - \frac{1-\mu_a^2}{n_o^2}},\ d\mu_o = \frac{\mu_a\,d\mu_a}{\sqrt{n_o^2-(1-\mu_a^2)}} \tag{5.86}$$

that transforms the integrals into the form

$$\int\limits_{-1}^{-\mu_t} Q_k^m(\mu_o)C^m(\tau,\mu_o)\,d\mu_o = \int\limits_{-1}^{0} Q_k^m(\mu_o)C^m(\tau,\mu_o)\frac{\mu_a\,d\mu_a}{\sqrt{n_o^2-(1-\mu_a^2)}},$$

$$\mu_o = -\sqrt{1-\frac{1-\mu_a^2}{n_o^2}}\,, \tag{5.87}$$

$$\int\limits_{\mu_t}^{1} Q_k^m(\mu_2)C^m(\tau,\mu_2)\,d\mu_2 = \int\limits_{0}^{1} Q_k^m(\mu_o)C^m(\tau,\mu_o)\frac{\mu_a\,d\mu_a}{\sqrt{n_o^2-(1-\mu_a^2)}},$$

$$\mu_o = \sqrt{1-\frac{1-\mu_a^2}{n_o^2}}\,. \tag{5.88}$$

It is easy to see that at the conversion to the discrete ordinates, perfect coincidence appears between the atmospheric $\vec{C}_+^a$, $\vec{C}_-^a$ and oceanic ordinates $\vec{C}_+^o$, $\vec{C}_-^o$. Thus, for the ocean, quadrature formula weights change as:

$$w_i^o = w_i^a\frac{\mu_i}{\sqrt{n_o^2-(1-\mu_i^2)}}\,, \tag{5.89}$$

where μ_i are the roots of the polynomial of degree N.

If we introduce vectors $\vec{C}_+^{ocn} = \left[\vec{C}_+^t;\ \vec{C}_+^o\right]$, $\vec{C}_-^{ocn} = \left[\vec{C}_-^o;\ \vec{C}_-^t\right]$ where square brackets designate a set of vectors that are combined in one with the summary dimension in accordance with the rules of the Matlab system, then all the relations (5.48)–(5.69) will be valid. Accordingly, for the introduced variables it is possible to write down the condition on the coupled interface of atmosphere–ocean:

$$\begin{bmatrix} \vec{C}^a_- \\ \vec{C}^{ocn}_+ \end{bmatrix} = \begin{bmatrix} \overleftrightarrow{R} & \overleftrightarrow{T}_{ao} \\ \overleftrightarrow{T}_{oa} & \overleftrightarrow{R}_{oo} \end{bmatrix} \begin{bmatrix} \vec{C}^a_+ \\ \vec{C}^{ocn}_- \end{bmatrix},$$

$$\overleftrightarrow{T}_{ao} \equiv [\, \overleftrightarrow{T} \;\; \overleftrightarrow{0} \,], \;\; \overleftrightarrow{T}_{oa} \equiv \begin{bmatrix} \overleftrightarrow{0} \\ \overleftrightarrow{T} \end{bmatrix}, \;\; \overleftrightarrow{R}_{oo} \equiv \begin{bmatrix} \overleftrightarrow{0} & \overleftrightarrow{1} \\ \overleftrightarrow{R} & \overleftrightarrow{0} \end{bmatrix}, \tag{5.90}$$

where $\overleftrightarrow{R}$, $\overleftrightarrow{T}$ are the Fresnel matrices.

The expression (5.90) allows us to use the matrix-operator method. It is not difficult to see that in the case of semi-infinite ocean with Fresnel border the radiance of the reflected diffuse radiation will get the form

$$\vec{C}_- = -\left[\overleftrightarrow{T} \;\; \overleftrightarrow{0} \right] \left(\overleftrightarrow{1} + \overleftrightarrow{u}^{-1}_{12} \overleftrightarrow{u}_{11} \overleftrightarrow{R}_{oo} \right)^{-1} \overleftrightarrow{u}^{-1}_{12} \vec{J}_- \,. \tag{5.91}$$

In Fig. 5.2 the comparison of the radiance calculation by the ordinary DOM with the subtraction only of the direct radiation and our approach with the removing of the anisotropic part of the solution made on the basis of MSH is shown. Here we use for the calculation the Henyey–Greenstein phase function:

$$p_{HG}(\gamma) = \frac{1 - g^2}{(1 + g^2 - 2g\,\cos\gamma)^{3/2}}\,, \tag{5.92}$$

that depends on only one parameter g – the average cosine of the scattering angle.

All the calculations in this chapter were executed under the condition that

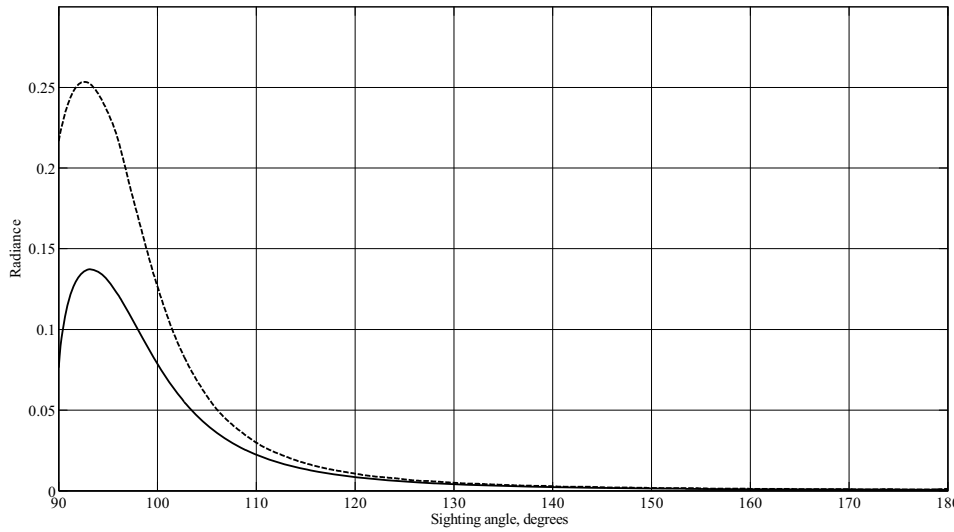
$$\int\limits_{(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0)\geq 0} L(\tau,\hat{\boldsymbol{l}})(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0)\, d\hat{\boldsymbol{l}} = 1.$$

The parameters of the medium: $\tau_0 = 1$, $\theta_0 = 60^\circ$, $\omega_{\text{w}} = 0.99$, $g = 0.98$ are used in this calculation. The calculation parameters are equal in the both the programs: $N = 81$, $M = 8$, $N_x = 451$, $K = N_x$. Of course, if $N = N_x$ and $M = 64$ are used in the ordinary DOM, its calculation results are identical to the suggested approach but the calculating time is increased by more than 100 times. It is easy to see that the algorithm used essentially increases the solution convergence that considerably raises the calculating speed.

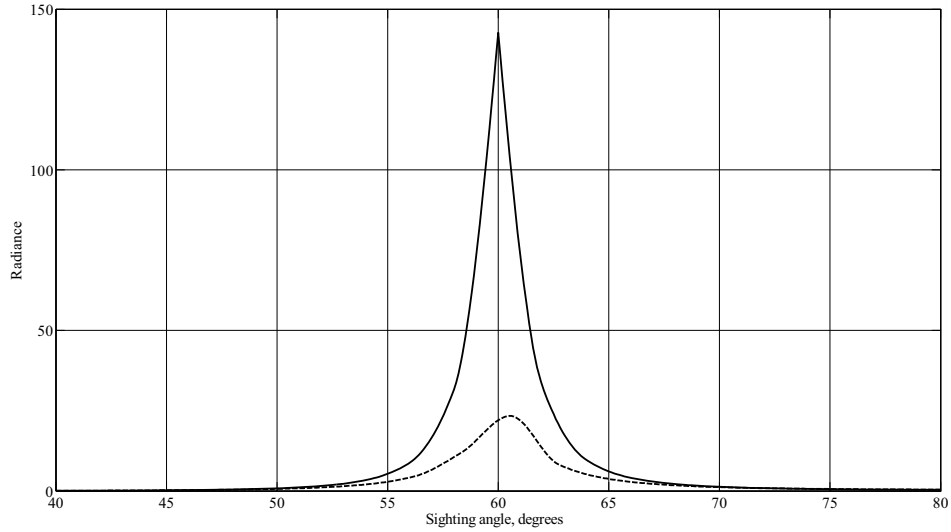
In Fig. 5.3 the numerical comparison of the angular distribution calculation of the radiation transmitted (a, c, e) and reflected (b, d, a) by the slab of two codes are shown: DOM (the proposed algorithm, a solid line) and the well-known code DISORT (a dotted line with black circles, ftp://climate1.gsfc.nasa.gov/wiscombe/Multiple_Scatt/). The charts are connected with the different types of scattering: Fig. 5.3(a,b), Water Haze L; (c,d), Water Cloud C1 (Deirmendjian, 1969); (e,f), ocean waters (Mobley et al., 2002). The physical model of Water Haze L and Water Cloud C1 by Deirmendjian (1969) is scattering on the ensemble of the spherical water drops with the modified gamma particle size distribution:

$$\frac{dN(r)}{dr} = ar^{\alpha-1}(\alpha - \gamma b r^\gamma)\, \exp(-br^\gamma)\,,$$

where $N(r)$ is the number of spherical particles with a radius smaller than r. We use the wavelength $\lambda = 0.7\,\mu$m, and the refractive index of water drops $n = 1.33$.

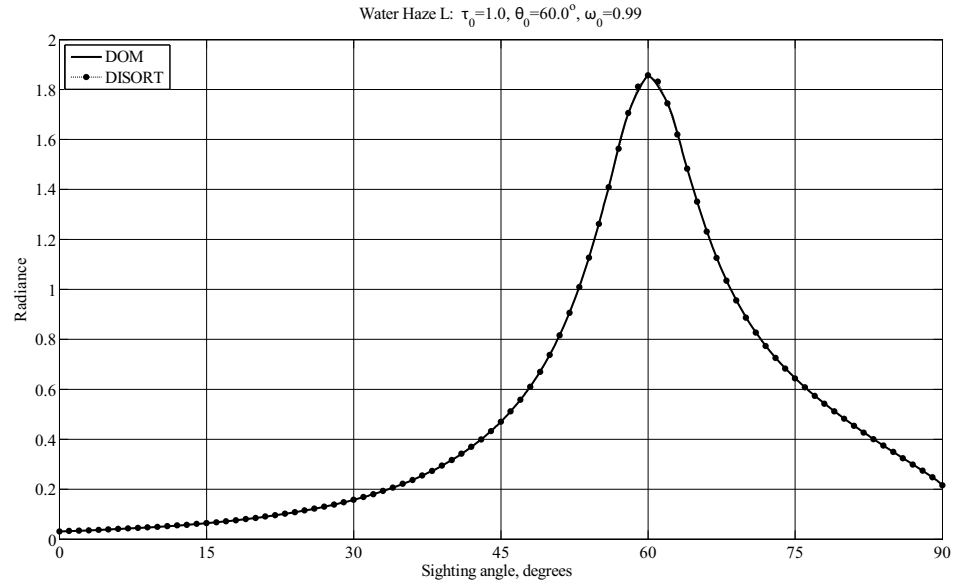


a)

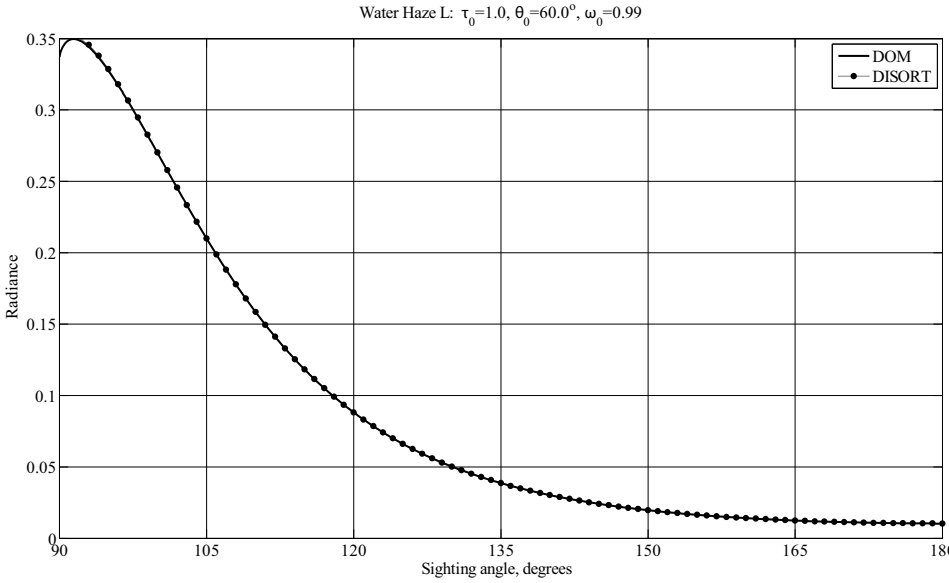


b)

Fig. 5.2. The comparison of the radiance calculation by the two different numerical methods: (a) reflected radiation, (b) transmitted radiation. The new suggested method is designated by the solid line, the 'ordinary' DOM is designated by the dashed line. The calculation parameters are equal in the both programs: $N = 81$, $M = 8$, $N_x = 451$, $K = N_x$. Of course, if $N = N_x$ and $M = 64$ are used in the ordinary DOM, its calculation results are identical to the suggested approach but the calculating time increases more than 100 times.

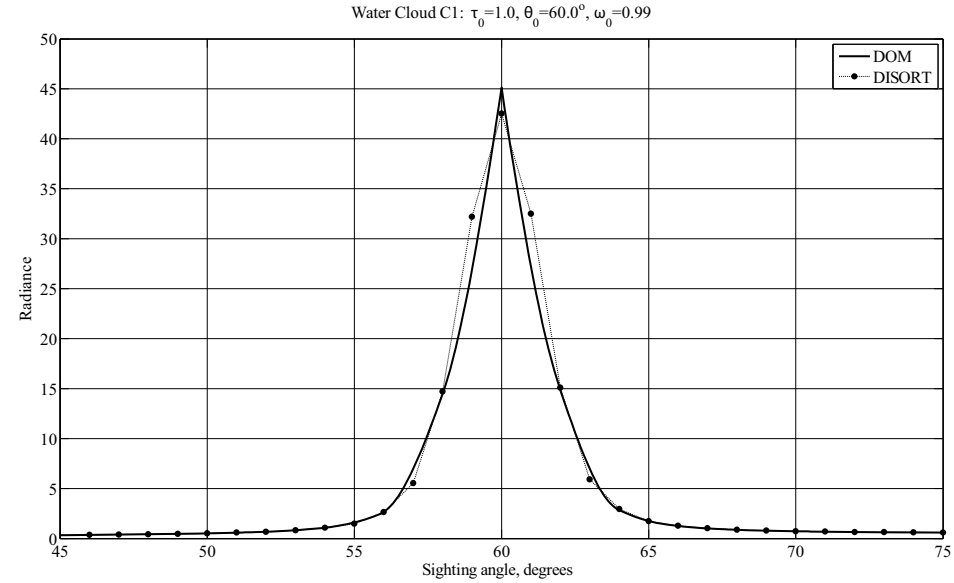


a)

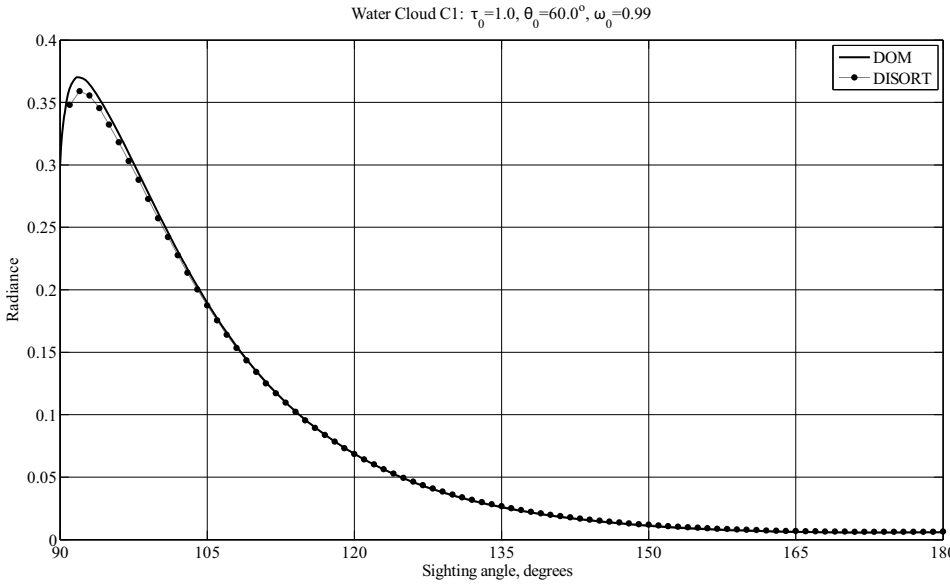


b)

Fig. 5.3. The numerical comparison of the angular distribution of the radiation reflected and transmitted by the slab was obtained using two codes: DOM and DISORT. The underlying surface albedo is $\rho = 0$. Figures (c) and (e) show the obvious advantage of the proposed method in the calculation of the forward sharp peak at the strong anisotropic scattering. It might be supposed that DISORT has a limitation in using the number of azimuth harmonics M, because its calculation time increases greatly, but we used the program as we downloaded it.

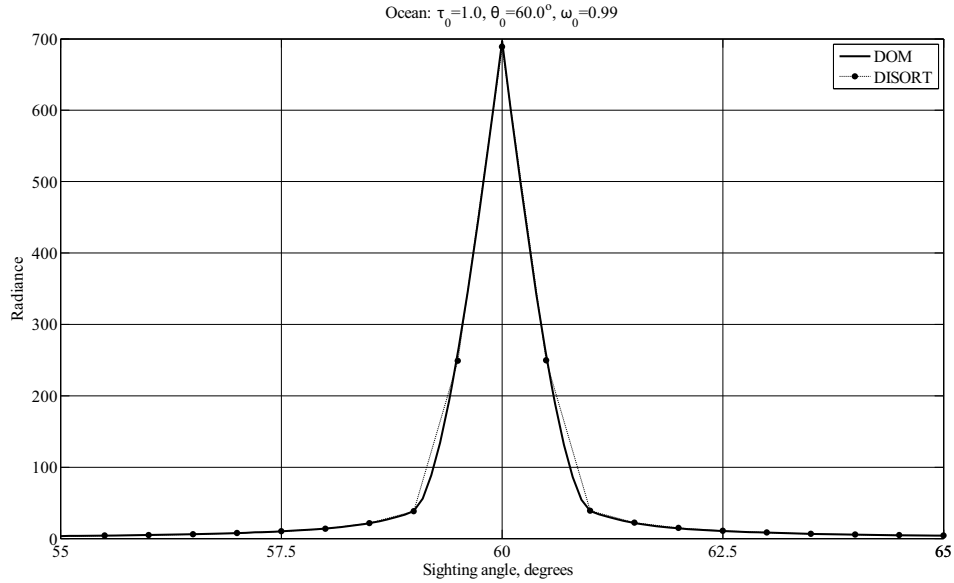


c)

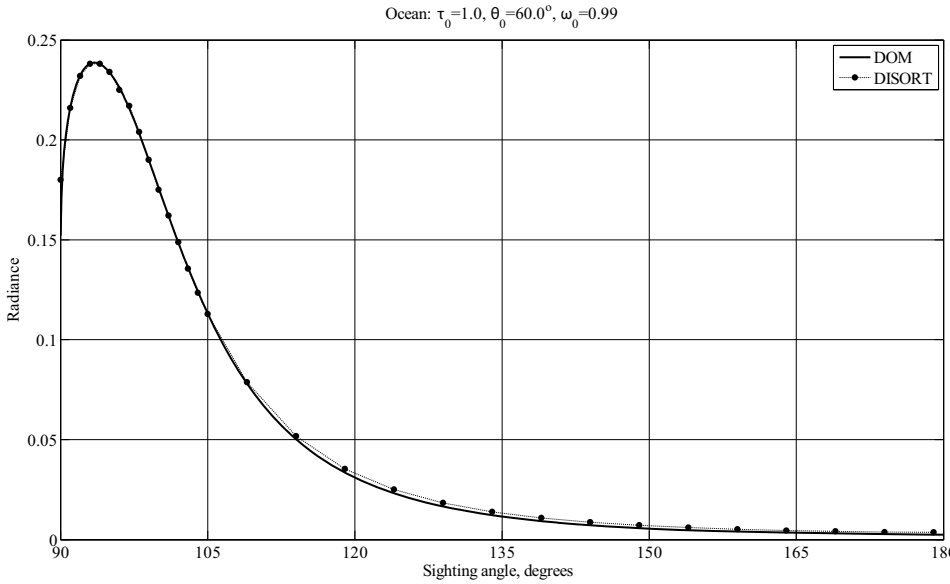


d)

Fig. 5.3. Continued.



e)



f)

Fig. 5.3. Continued.

The parameters of the particle size distribution were: $a = 4.9757 \cdot 10^6$, $\alpha = 2$, $b = 15.1186$, $\gamma = 0.5$ for Water Haze L and $a = 2.373$, $\alpha = 6$, $b = 1.5$, $\gamma = 1$ for Water Cloud C1.

The calculation parameter $N = 152$ (the number of ordinates) was equal for the both codes. The calculation time (PC, CPU Intel Core2Quad Q9550 @ 2.83 GHz, RAM 4Gbyte, Windows XP) for DOM (Matlab 2009a): Water Haze L, 0.1 s; Water Cloud C1, 0.13 s; Ocean, 1.3 s; for DISORT (Intel Visual Fortran 11): Water Haze L, 14.7 s; Water Cloud C1, 57.4 s; Ocean, 487.4 s. It is easy to see that for Water Haze L the both algorithms are equivalent in accuracy, but for the more anisotropic scattering the code DISORT demands essentially more calculation time for the same accuracy.

Of course, the comparison obcodes is very relative, as they are realized in different programming systems. However, it is definitely possible to state that the computational results of the proposed algorithm are more stable and smooth functions, especially in the case of strong anisotropic scattering.

5.7 Generalization for the vectorial case of polarized radiation

The spatial distribution of the Stokes vector $\vec{L}(\tau, \hat{\boldsymbol{l}}) = \left[\begin{array}{cccc} I & Q & U & V \end{array}\right]^T$ describes all the optical properties of a light beam (Rozenberg, 1977). The components of the Stokes vector can be measured experimentally and determined by the responses J_i of four optical detectors with different polarization filters (Rozenberg, 1977)

$$I = 2J_1, \quad Q = 2(J_1 - J_2), \quad U = 2(J_3 - J_1), \quad V = 2(J_4 - J_1)\,, \tag{5.93}$$

where J_1 is the neutral filter with the transmittance $\tau = 0.5$; J_2 is the polarization analyzer with the optical axis and the direction of radiation propagation that determine the reference plane; J_3 is the polarization analyzer with axis angularly 45° to the reference plane; J_4 is the complex filter composed of quarter-wave plate and the polarization analyzer with axis angularly 45° to the reference plane.

Then the response of the arbitrary detector can be expressed in the following form

$$J = 0.5\tau(I + Q\,\cos 2\varphi + \sin 2\varphi(U\cos\delta + V\sin\delta)) = 0.5\tau\left\{1, \vec{D}\right\}\vec{L}\,, \tag{5.94}$$

where τ is the optical system transmittance, $\{1,\ \vec{D}\}$ is the 4-D vector; φ is the angle between the polarization analyzer and the reference plane; δ is the phase shift of the phase plate; $\vec{D} = \{\cos 2\varphi, \sin 2\varphi\cos\delta, \sin 2\varphi \sin 2\delta\}$ is the vector of polarization features of the receiving system.

Thus, the Stokes vector is the generalization of the radiance in the case of polarization. This Stokes vector submits to the boundary problem for the vectorial RTE (VRTE) and a slab irradiated by a plane unidirectional source (Chandrasekhar, 1950) similar to (5.5)

$$\begin{cases} \mu\dfrac{\partial\vec{L}(\tau,\hat{\boldsymbol{l}})}{\partial\tau} + \vec{L}(\tau,\hat{\boldsymbol{l}}) = \dfrac{\omega_0}{4\pi}\oint \vec{S}(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')\vec{L}(\tau,\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}'; \\ \vec{L}(0,\hat{\boldsymbol{l}})\Big|_{\mu\geq 0} = \vec{L}_0\delta(\hat{\boldsymbol{l}} - \hat{\boldsymbol{l}}_0);\ \vec{L}(\tau_0,\hat{\boldsymbol{l}})\Big|_{\mu\leq 0} = 0, \end{cases} \tag{5.95}$$

where the local transformation matrix $\overleftrightarrow{S}(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')$ has the following form including the rotation of a reference plane while the scattering act occurs

$$\overleftrightarrow{S}(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')=\overleftrightarrow{R}(\hat{\boldsymbol{l}}\times\hat{\boldsymbol{l}}'\to\hat{\boldsymbol{l}}_0\times\hat{\boldsymbol{z}})\overleftrightarrow{p}(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')\overleftrightarrow{R}(\hat{\boldsymbol{l}}_0\times\hat{\boldsymbol{z}}\to\hat{\boldsymbol{l}}\times\hat{\boldsymbol{l}}')=\overleftrightarrow{R}(\chi)\overleftrightarrow{p}(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')\overleftrightarrow{R}(\chi')\,, \qquad (5.96)$$

$\overleftrightarrow{R}(\cdot)$ is the rotator, χ is a dihedral angle between the $\hat{\boldsymbol{l}}\times\hat{\boldsymbol{l}}'$ and $\hat{\boldsymbol{l}}_0\times\hat{\boldsymbol{z}}$ planes, $\overleftrightarrow{p}(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')$ is a scattering matrix, $\vec{L}_0$ is the Stokes vector for the top boundary condition.

As in the scalar case, one of the main difficulties in the VRTE solution is the scattering integral evaluation. However, in the vectorial case, there is still another problem: the reference planes of the incident and scattering rays as well as the scattering plane do not coincide (Kuščer and Ribarič, 1959). Therefore, it is necessary to use the rotator $\overleftrightarrow{R}$ that disturbs the transformation symmetry of the different Stokes parameters and makes it impossible to use the addition theorem for the surface harmonics.

For the case of polarized radiation transfer, the well-known circular basis (Kuščer and Ribarič, 1959) is the only form to evaluate the scattering integral for any type of scattering matrix. This form is called the CP-representation (Circular Polarization) and can be obtained from the real-number energetic SP-representation (Stokes Polarization) using the known matrix transformation containing complex numbers. Thus CP-representation is the complex basis but it makes it possible to diagonalize the rotation matrix in (5.96) as

$$\begin{aligned}
\overleftrightarrow{R}_{CP}(\chi) &= \overleftrightarrow{T}_S\overleftrightarrow{R}_{SP}(\chi)\overleftrightarrow{T}_S^{-1} \\
&= \frac{1}{2}\begin{bmatrix} 0 & 1 & -i & 0 \\ 1 & 0 & 0 & -1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & i & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\chi & \sin 2\chi & 0 \\ 0 & -\sin 2\chi & \cos 2\chi & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ i & 0 & 0 & -i \\ 0 & -1 & 1 & 0 \end{bmatrix} \\
&= \operatorname{diag}\left[\ \mathrm{e}^{i2\chi}\ \ 1\ \ 1\ \ \mathrm{e}^{-i2\chi}\ \right] = \operatorname{diag}\left[\ \mathrm{e}^{+i2\chi}\ \ \mathrm{e}^{+i0\chi}\ \ \mathrm{e}^{-i0\chi}\ \ \mathrm{e}^{-i2\chi}\ \right]. \qquad (5.97)
\end{aligned}$$

In the complex circular basis, the spatial spectrum of the desired Stokes vector and the scattering matrix are always written as the general spherical function (GSF) (Gelfand et al., 1963) expansion

$$\begin{aligned}
\vec{L}(\tau,\hat{\boldsymbol{l}}) &= \sum_{m=-\infty}^{\infty}\sum_{k=0}^{\infty}\frac{2k+1}{4\pi}\overleftrightarrow{Y}_m^k(\mu)\vec{f}_m^k(\tau)\,\exp(im\varphi)\,, \\
\left[\overleftrightarrow{p}(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')\right]_{r,s} &= \sum_{k=0}^{\infty}(2k+1)x_{r,s}^k P_{r,s}^k(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')\,, \qquad (5.98)
\end{aligned}$$

with indices r and s running through all $n = +2, +0, -2, -0$ simultaneously and independently in accordance with the exponential powers in (5.97) with the generalized Legendre polynomials $P_{m,n}^k$ having the matrix form (Ustinov, 1988):

$$\overleftrightarrow{Y}_m^k(\mu) = \operatorname{diag}\left[\ P_{m,+2}^k(\mu);\ \ P_{m,+0}^k(\mu);\ \ P_{m,-0}^k(\mu);\ \ P_{m,-2}^k(\mu)\ \right]. \qquad (5.99)$$

We use the following definition for the GSF (Gelfand et al., 1963)

$$P^k_{m,n}(\mu) = \frac{(-1)^{k-m} i^{n-m}}{2^k (k-m)!} \sqrt{\frac{(k-m)!(k+n)!}{(k+m)!(k-n)!}} \cdot (1-\mu)^{-\frac{n-m}{2}} (1+\mu)^{-\frac{n+m}{2}} \frac{d^{k-n}}{d\mu^{k-n}} \left[(1-\mu)^{k-m}(1+\mu)^{k+m}\right] .$$

The special form of the addition theorem is correct for GSF (Gelfand et al., 1963):

$$\mathrm{e}^{-im\chi} P^l_{mn}(\hat{\boldsymbol{l}} \cdot \hat{\boldsymbol{l}}') \, \mathrm{e}^{-in\chi'} = \sum_{q=-l}^{l} (-1)^q P^l_{mq}(\hat{\boldsymbol{l}} \cdot \hat{\boldsymbol{z}}) P^l_{qn}(\hat{\boldsymbol{z}} \cdot \hat{\boldsymbol{l}}') \, \mathrm{e}^{iq(\varphi-\varphi')} . \qquad (5.100)$$

The diagonalization of the rotation matrix (5.97) of the VRTE together with (5.98), the addition theorem (5.100) and some recurrent formulas for the GSF allow us to derive a system of the vectorial differential equations for the spherical harmonics (Ustinov, 1988) or the discrete ordinates (Siewert, 2000) methods. The solution of the system becomes possible only after addition of some boundary conditions – Mark's form for the boundary conditions (Mark, 1947) that is widely used. To improve the conditionality of the system matrix and eliminate the solution oscillation one has to apply the scaling transformation (Karp et al., 1980) for the system obtained.

In order to generalize MSH for the vectorial case of polarized radiation we will change the reference frame of the sighting direction counting from $\hat{\boldsymbol{l}}(\mu, \varphi)$to the reference frame $\hat{\boldsymbol{l}}(\nu, \psi)$, where ν is counted from the top of the slab irradiance direction $\hat{\boldsymbol{l}}_0$, and ψ lies in the perpendicular to the $\hat{\boldsymbol{l}}_0$ plane. The mutual relation for this systems is given after (5.11). This relation gives the transformation of the differential the item of the VRTE: both the item with a sighting angle cosine ν and an item with a sighting angle sine $\sqrt{1-\nu^2}$ originate from the described frame of the reference change. We convert both of the above-mentioned items using some formulas derived from (Gelfand et al., 1963) and expand the scattering integral using the method described in (Kuščer and Ribarič, 1959; Ustinov, 1988). Finally, we obtain a system of equations for the desired coefficients in CP-representation in the following form

$$\begin{aligned} &\frac{1}{2k+1} \frac{\partial}{\partial \tau} \left\{ \mu_0 \left[\overleftrightarrow{A}^{k+1}_m \vec{f}^{k+1}_m(\tau) + \overleftrightarrow{B}^{k}_m \vec{f}^{k}_m(\tau) + \overleftrightarrow{A}^{k}_m \vec{f}^{k-1}_m(\tau) \right] \right. \\ &+ \frac{i}{2} \sqrt{1-\mu_0^2} \left[\overleftrightarrow{h}_1 \vec{f}^{k-1}_{m-1}(\tau) + \overleftrightarrow{h}_2 \vec{f}^{k}_{m-1}(\tau) + \overleftrightarrow{h}_3 \vec{f}^{k+1}_{m-1}(\tau) \right. \\ &\left. \left. + \overleftrightarrow{h}_4 \vec{f}^{k-1}_{m+1}(\tau) + \overleftrightarrow{h}_5 \vec{f}^{k}_{m+1}(\tau) + \overleftrightarrow{h}_6 \vec{f}^{k+1}_{m+1}(\tau) \right] \right\} \\ &+ (\overleftrightarrow{1} - \omega_0 \overleftrightarrow{p}_k) \vec{f}^{k}_m(\tau) = \vec{0}, \end{aligned} \qquad (5.101)$$

where (n runs through $+2$, $+0$, -2, -0, and δ is the Kronecker symbol)

$$\left[\overleftrightarrow{A}^{k}_m\right]_{r,s} = \frac{1}{k} \sqrt{(k^2-m^2)(k^2-s^2)} \delta_{r,s}; \left[\overleftrightarrow{B}^{k}_m\right]_{r,s} = \frac{ms}{k(k+1)} (2k+1) \delta_{r,s}$$

$$\overleftrightarrow{h}_1 = \operatorname{diag}\left[\frac{\sqrt{(k+n+1)(k-n+1)(k+m)(k+m+1)}}{(k+1)}\right],$$

$$\overleftrightarrow{h}_2 = -\operatorname{diag}\left[\frac{n(2k+1)\sqrt{(k+m)(k-m+1)}}{k(k+1)}\right]$$

$$\overleftrightarrow{h}_3 = -\operatorname{diag}\left[\frac{\sqrt{(k+n)(k-n)(k-m)(k-m+1)}}{k}\right];$$

$$\overleftrightarrow{h}_6 = -\operatorname{diag}\left[\frac{\sqrt{(k+n)(k-n)(k+m)(k+m+1)}}{k}\right]$$

$$\overleftrightarrow{h}_5 = \operatorname{diag}\left[\frac{n(2k+1)\sqrt{(k+m+1)\,(k-m)}}{k\,(k+1)}\right],$$

$$\overleftrightarrow{h}_4 = \operatorname{diag}\left[\frac{\sqrt{(k+n+1)\,(k-n+1)\,(k-m)\,(k-m+1)}}{(k+1)}\right]$$

The significant anisotropy of the phase function and the feature of a light field inside the slab (δ-singularity due to the sight direction) allow us to assume the scalar case that the light field is an essentially elongated function along the direction of the incident light. Therefore, its spatial spectrum coefficients (5.98) are the smooth functions via the GSF order number k. If we assume a continuous dependence of $\vec{f}^m$ instead of the discrete one with k as an argument we will be able to cut the Taylor expansion of this function to two terms due to its smoothness. In the vectorial form, as we did with expression (5.14), we can write as:

$$\vec{f}^m(\tau, k \pm 1) \approx \vec{f}^m(\tau, k) \pm \frac{\partial \vec{f}^m(\tau, k)}{\partial k}. \tag{5.102}$$

We note here that more than two terms in (5.102) lead to significant analytical evaluation problems. In addition to that, the anisotropy of the light field is much greater than its asymmetry. This fact can be expressed in a simple way in the chosen frame of reference as $k \gg m > n$ (k belongs to the ν-expansion, m belongs to the ψ-expansion, n is the polarization index). The above-mentioned inequality and the reassignment for κ (5.14) allow us to simplify matrices in (5.101) and to write them as

$$\overleftrightarrow{A}^k_m \approx k\overleftrightarrow{1};\ \overleftrightarrow{B}^k_m = \frac{2\,(2k+1)\,m}{k\,(k+1)}\overleftrightarrow{O};\ \overleftrightarrow{O} = \operatorname{diag}\begin{bmatrix}1 & 0 & 0 & -1\end{bmatrix};\ \overleftrightarrow{h}_1 \approx \left(\kappa + \frac{(2k+1)m}{2\kappa}\right)\overleftrightarrow{1};$$

$$\overleftrightarrow{h}_2 \approx -\frac{2\,(2k+1)}{\kappa}\overleftrightarrow{O};\ \overleftrightarrow{h}_3 \approx -\left(\kappa - \frac{(2k+1)m}{2\kappa}\right)\overleftrightarrow{1};\ \overleftrightarrow{h}_4 = -\overleftrightarrow{h}_3;\ \overleftrightarrow{h}_5 = -\overleftrightarrow{h}_2;\ \overleftrightarrow{h}_6 = -\overleftrightarrow{h}_1,$$

Along with (5.102), we convert (5.101) into the expression:

$$\mu_0 \frac{\partial \vec{f}^m(\tau,\kappa)}{\partial \tau} + \frac{\partial}{\partial \tau}\left\{\mu_0\left(\frac{\partial \vec{f}^m(\tau,\kappa)}{2\kappa\partial\kappa} + \frac{2m}{\kappa^2}\overleftrightarrow{O}\vec{f}^k_m(\tau,\kappa)\right) + \right.$$
$$+\frac{\sqrt{1-\mu_0^2}}{2}\left[-\left(\frac{\partial \vec{f}^{m-1}(\tau,\kappa)}{\partial\kappa} + \frac{\partial \vec{f}^{m+1}(\tau,\kappa)}{\partial\kappa}\right)\right.$$
$$\left.\left.+\frac{m}{\kappa}\left(\vec{f}^{m-1}(\tau,\kappa) - \vec{f}^{m+1}(\tau,\kappa)\right) + \frac{2}{\kappa}\overleftrightarrow{O}(\vec{f}^{m+1}(\tau,\kappa) - \vec{f}^{m-1}(\tau,\kappa))\right]\right\}$$
$$= -(\overleftrightarrow{1} - \omega_0\overleftrightarrow{p}_k)\vec{f}^m(\tau,\kappa)\,. \tag{5.103}$$

We use the method of scalar small angle modification of spherical harmonics (5.16) to find the solution of the expression (5.103). To do this, we introduce a special function $\vec{\omega} = \vec{\omega}(\tau,\psi,k)$ such as its Fourier-spectrum is the desired coefficients $\vec{f}^m$:

$$\vec{w}(\tau,\psi,\kappa) = \sum_{m=-\infty}^{\infty} \vec{f}^m(\tau,\kappa)\,\exp(im\psi);\;\; \vec{f}^m(\tau,\kappa) = \frac{1}{2\pi}\int_0^{2\pi} \vec{w}(\tau,\psi,\kappa)\,\exp(-im\psi)\,d\psi \tag{5.104}$$

Some properties of (5.104) can be proved according to (5.17). The expressions (5.104) with the directional derivation of the projection of the incident direction to the plane of the ψ-angle can be written as

$$(\boldsymbol{l}_{0\perp}, \nabla_\kappa)\vec{w}(\tau,\psi,\kappa) = \sqrt{1-\mu_0^2}\left(\cos\psi\frac{\partial}{\partial\kappa}\vec{w}(\tau,\psi,\kappa) + \frac{\sin\psi}{\kappa}\frac{\partial}{\partial\psi}\vec{w}(\tau,\psi,\kappa)\right).$$

Now we can transform (5.103) into a simple equation

$$\frac{\partial}{\partial\tau}\left[\mu_0 - i(\hat{\boldsymbol{l}}_{0\perp}, \nabla_\kappa)\right]\vec{w}(\tau,\psi,\kappa) = -(\overleftrightarrow{1} - \omega_0\overleftrightarrow{p}_k)\vec{w}(\tau,\psi,\kappa)\,. \tag{5.105}$$

Since by equation (5.105) we seek only the anisotropic part of the complete solution and then will solve the boundary value problem for the regular part, one of the requirements to the MSH solution is analytical simplicity. But the solution of equation (5.105) is not very simple. Therefore, we go from (5.105) to the simpler equation that follows from (5.101) if the terms containing $\sqrt{1-\mu_0^2}$ are neglected, or the equivalent form of the approximation $\mu_0 \to 1$ is assumed. In this case, we get from (5.101) a simple vectorial differential equation just as we get from (5.19):

$$\mu_0\frac{\partial \vec{f}^m(\tau,k)}{\partial\tau} + (\overleftrightarrow{1} - \omega_0\overleftrightarrow{p}_k)\vec{f}^m(\tau,k) = \mathbf{0}\,, \tag{5.106}$$

that for the depth-homogeneous media has the solution in the following simple form

$$\vec{f}^m(\tau,k) = \exp(-(\overleftrightarrow{1} - \omega_0\overleftrightarrow{p}_k)\tau\Big/\mu_0)\vec{f}^m(0,k)\,. \tag{5.107}$$

In CP-representation we have

$$\vec{L}^{CP}_{MSH}(\tau,\nu,\psi) = \sum_{m=-2,0,2}\sum_{k=0}^{\infty}\frac{2k+1}{4\pi}\overleftrightarrow{Y}^m_k(\nu)$$
$$\cdot\,\exp\left\{-(\overleftrightarrow{1} - \omega_0\overleftrightarrow{p}_k)\tau\Big/\mu_0\right\}\exp(im\psi)\vec{f}^m_k(0)\,. \tag{5.108}$$

with the boundary conditions in CP

$$\begin{aligned}\vec{f}^{-2}(0) &= \left[\,0;\;\,0;\;\,0;\;\,p\,\right]^T;\\ \vec{f}^{0}(0) &= \left[\,0;\;\,1-q;\;\,1-q;\;\,0\,\right]^T;\\ \vec{f}^{+2}(0) &= \left[\,p;\;\,0;\;\,0;\;\,0\,\right]^T,\end{aligned}$$

where p is the degree of linear polarization, q is the ellipticity.

In SP we have

$$\begin{aligned}\vec{\vec{Z}}_k(\tau) &= \vec{\vec{T}}_{SC}\exp\left\{-(\vec{\vec{1}}-\omega_0\vec{\vec{p}}_k)\tau\Big/\mu_0\right\}\vec{\vec{T}}_{CS}\\ &= \exp\left\{-(\vec{\vec{1}}-\omega_0\vec{\vec{T}}_{SC}\vec{\vec{p}}_k\vec{\vec{T}}_{CS})\tau\Big/\mu_0\right\}\\ &= \exp\left\{-(\vec{\vec{1}}-\omega_0\vec{\vec{\chi}}_k)\tau\Big/\mu_0\right\},\end{aligned}\tag{5.109}$$

and the MSH itself has the form

$$\begin{aligned}&\vec{L}^{SP}_{MSH}(\tau,\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0)\\ &=\sum_{k=0}^{\infty}\frac{2k+1}{2\pi}\left\{\vec{\vec{P}}_R^{\,k,0}\vec{\vec{Z}}_k(\tau)\vec{f}^{0}(0)-\vec{\vec{P}}_R^{\,k,2}(\nu)\vec{\vec{Z}}_k(\tau)\vec{f}^{2}(0)\;-\vec{\vec{P}}_I^{\,k,2}(\nu)\vec{\vec{Z}}_k(\tau)\vec{f}^{-2}(0)\right\},\end{aligned}\tag{5.110}$$

with the vectors

$$\begin{aligned}\vec{f}^{-2}(0) &= \left[\,0\;\;p\sin 2\varphi'\;\;p\cos 2\varphi'\;\;0\,\right]^T;\\ \vec{f}^{0}(0) &= \left[\,1\;\;0\;\;0\;\;q\,\right]^T;\\ \vec{f}^{+2}(0) &= \left[\,0\;\;p\cos 2\varphi'\;\;p\sin 2\varphi'\;\;0\,\right]^T,\end{aligned}$$

where $\varphi\prime=\psi-\varphi_0$ – azimuth deviation.

Matrix polynomials are

$$\begin{aligned}\vec{\vec{P}}_R^{\,k,m}(\mu) &= \begin{bmatrix}Q_k^m(\mu) & 0 & 0 & 0\\ 0 & R_k^m(\mu) & 0 & 0\\ 0 & 0 & R_k^m(\mu) & 0\\ 0 & 0 & 0 & Q_k^m(\mu)\end{bmatrix};\\ \vec{\vec{P}}_I^{\,k,m}(\mu) &= \begin{bmatrix}0 & 0 & 0 & 0\\ 0 & 0 & T_k^m(\mu) & 0\\ 0 & -T_k^m(\mu) & 0 & 0\\ 0 & 0 & 0 & 0\end{bmatrix},\end{aligned}\tag{5.111}$$

with the notation $R_2^k(\nu)=0.5i^m(P^k_{2,2}(\nu)+P^k_{2,-2}(\nu))\quad T_2^k(\nu)=0.5i^m(P^k_{2,2}(\nu)-P^k_{2,-2}(\nu))$ (Siewert, 2000). SP-representation of (5.110) could have a different form with the analytically evaluated exponential matrix (Astakhov et al., 1994). However, the form (5.110) with direct computation is more useful.

5.8 Evaluation of the vectorial regular part

Following the scalar case, we represent the complete solution of (5.36) in the following form (Budak and Korkin, 2008a,b,c)

$$\vec{L}(\tau,\hat{\boldsymbol{l}}) = \vec{L}_{MSH}(\tau,\hat{\boldsymbol{l}}) + \tilde{L}(\tau,\hat{\boldsymbol{l}}). \tag{5.112}$$

The expansion (5.112) allows us to reduce the complete VRTE boundary problem to the one for the regular part only

$$\begin{cases} \mu\dfrac{\partial\tilde{L}(\tau,\hat{\boldsymbol{l}})}{\partial\tau} + \tilde{L}(\tau,\hat{\boldsymbol{l}}) = \dfrac{\omega_0}{4\pi}\oint \vec{S}(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')\tilde{L}(\tau,\hat{\boldsymbol{l}}')\,d\,\hat{\boldsymbol{l}}' + \vec{\Delta}(\tau,\hat{\boldsymbol{l}}); \\ \tilde{L}(0,\ \mu>0,\ \varphi) = \mathbf{0};\ \ \tilde{L}(\tau_0,\ \mu\le 0,\ \varphi) = -\vec{L}_{MSH}(\tau_0,\ \mu\le 0,\ \varphi), \end{cases} \tag{5.113}$$

where $\vec{\Delta}(\tau,\hat{\boldsymbol{l}})$ is the source function built upon the MSH. We substitute (5.112) in the VRTE (5.113) and obtain the following expression for the source function

$$\vec{\Delta}(\tau,\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0) = -\mu\frac{\partial}{\partial\tau}\vec{L}_{MSH}(\tau,\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0) - \vec{L}_{MSH}(\tau,\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0) + \frac{\omega_0}{4\pi}\oint \vec{S}(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')\vec{L}_{MSH}(\tau,\hat{\boldsymbol{l}}',\hat{\boldsymbol{l}}_0)\,d\hat{\boldsymbol{l}}'. \tag{5.114}$$

We especially note here that the source function upon the MSH does not change the form of the boundary problem as the classical approach does. Hence, the evaluation of the smooth part could be carried out with any of the modern well-developed methods of VRTE solution: discrete ordinates method (DOM), spherical harmonics method, Monte Carlo simulation (Marchuk et al., 1980) and so on. In the present chapter, we use DOM (modified by double Gaussian quadratures (Sykes, 1951)) for both scalar and vectorial problems. To solve (5.113) for this method it is very convenient to include boundary conditions with complicated geometry in the Mark form.

We evaluate the scattering integral of the VRTE in CP-basis that allows us to diagonalize the rotation matrix. While evaluating the scattering integral we follow the well-known routine: we substitute (5.98) in the scattering integral (5.113), and we use the addition theorem and the orthogonal properties of the GSF. After expanding the scattering integral we must use the scaling transformation to make the solution well-conditioned. For this purpose we use the universal transformation from the complex-numbered CP- to the real-numbered SP-basis. Further on we will illustrate the main steps of the abovementioned algorithm.

We denote the SP-to-CP transformation matrix with $\overleftrightarrow{T}_{CS}$ and with $\overleftrightarrow{T}_{SC}^{-1} = \overleftrightarrow{T}_{CS}$ - the universal matrix (5.97). Let's represent the addition theorem in the following form

$$\overleftrightarrow{R}(\chi)\overleftrightarrow{Y}_2^k(\hat{\boldsymbol{l}}\cdot\hat{\boldsymbol{l}}')\overleftrightarrow{R}(\chi') = \sum_{m=-k}^{k}(-1)^m\overleftrightarrow{Y}_{2m}^k(\hat{\boldsymbol{l}}\cdot\hat{\boldsymbol{z}})\overleftrightarrow{Y}_{m2}^k(\hat{\boldsymbol{z}}\cdot\hat{\boldsymbol{l}}')\,\mathrm{e}^{im(\varphi-\varphi')}\,,$$

and the local transformation matrix (5.96) as

$$\overleftrightarrow{R}_{CP}(\chi)\left[p_{rs}^k P_{r,s}^k(\hat{\boldsymbol{l}}\hat{\boldsymbol{l}}')\right]\overleftrightarrow{R}_{CP}(\chi') = \sum_{m=-k}^{k}(-1)^m\,\mathrm{e}^{im(\varphi-\varphi')}\overleftrightarrow{Y}_m^k(\mu)\overleftrightarrow{p}_k\overleftrightarrow{Y}_m^k(\mu')\,,$$

where $\overleftrightarrow{p}_k = \left[x^k_{rs}\right]$. Using this we rewrite the scattering integral as

$$\begin{aligned}\vec{I}_{SP} &= \frac{\omega_0}{4\pi}\sum_{k=0}^{\infty}(2k+1)\sum_{m=-k}^{k}(-1)^m\overleftrightarrow{T}_{SC}\overleftrightarrow{Y}^k_m(\mu)\overleftrightarrow{p}_k \\ &\cdot\oint \mathrm{e}^{im(\varphi-\varphi')}\overleftrightarrow{Y}^k_m(\mu')\overleftrightarrow{T}_{CS}\tilde{L}_{SP}(\tau,\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}'. \end{aligned} \tag{5.115}$$

Equation (5.115) can be written in more symmetric form by getting the transformation matrix:

$$\begin{aligned}\vec{I}_{SP} &= \frac{\omega_0}{4\pi}\sum_{k=0}^{\infty}(2k+1)\sum_{m=-k}^{k}(-1)^m\overleftrightarrow{T}_{SC}\overleftrightarrow{Y}^k_m(\mu)\overleftrightarrow{T}_{CS}\overleftrightarrow{T}_{SC}\overleftrightarrow{p}_k\overleftrightarrow{T}_{CS} \\ &\cdot\oint \mathrm{e}^{im(\varphi-\varphi')}\overleftrightarrow{T}_{SC}\overleftrightarrow{Y}^k_m(\mu')\overleftrightarrow{T}_{CS}\tilde{L}_{SP}(\tau,\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}'. \end{aligned}$$

We define new matrices: $\overleftrightarrow{\chi}_k \equiv \overleftrightarrow{T}_{SC}\overleftrightarrow{p}_k\overleftrightarrow{T}_{CS}$, $\overleftrightarrow{P}^k_m(\mu) \equiv \overleftrightarrow{T}_{SC}\overleftrightarrow{Y}^k_m(\mu)\overleftrightarrow{T}_{CS}$, and transform the scattering integral to the following form

$$\vec{I}_{SP} = \frac{\Lambda}{4\pi}\oint\left(\sum_{k=0}^{\infty}(2k+1)\sum_{m=-k}^{k}\exp(-im(\varphi-\varphi'))\overleftrightarrow{P}^k_m(\mu)\overleftrightarrow{\chi}_k\overleftrightarrow{P}^k_m(\mu)\right)\tilde{L}(\tau,\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}'. \tag{5.116}$$

Then we represent matrix polynomials as the sum of their real and imaginary parts (5.111)

$$\overleftrightarrow{P}^k_m(\mu) = \begin{bmatrix} Q^m_k(\mu) & 0 & 0 & 0 \\ 0 & R^m_k(\mu) & -iT^m_k(\mu) & 0 \\ 0 & iT^m_k(\mu) & R^m_k(\mu) & 0 \\ 0 & 0 & 0 & Q^m_k(\mu) \end{bmatrix} = \overleftrightarrow{P}^{k,m}_R(\mu) + i\,\overleftrightarrow{P}^{k,m}_I(\mu). \tag{5.117}$$

We made the subsequent transformations of the scattering integral (5.116) after combining all real and imaginary parts taking into account (5.117) and the Euler formula for the complex azimuth-dependent exponent in (5.116). Then we use the even properties of trigonometric functions. After some algebraic transformations we substitute equation (5.114) by

$$\begin{aligned}\vec{I}_{SP} &= \frac{\omega_0}{4\pi}\oint\sum_{k=0}^{\infty}(2k+1)\sum_{m=0}^{k}(2-\delta_{0,m})\times \\ &\times\left[\overleftrightarrow{C}^m_k(\mu,\mu')\cos m(\varphi-\varphi') + \overleftrightarrow{S}^m_k(\mu,\mu')\sin m(\varphi-\varphi')\right]\tilde{L}(\tau,\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}', \end{aligned} \tag{5.118}$$

where

$$\begin{aligned}\overleftrightarrow{C}^m_k(\mu,\mu') &= \overleftrightarrow{P}^{k,m}_R(\mu)\overleftrightarrow{\chi}_k\overleftrightarrow{P}^{k,m}_R(\mu') - \overleftrightarrow{P}^{k,m}_I(\mu)\overleftrightarrow{\chi}_k\overleftrightarrow{P}^{k,m}_I(\mu'), \\ \overleftrightarrow{S}^m_k(\mu,\mu') &= \overleftrightarrow{P}^{k,m}_I(\mu)\overleftrightarrow{\chi}_k\overleftrightarrow{P}^{k,m}_I(\mu') + \overleftrightarrow{P}^{k,m}_R(\mu)\overleftrightarrow{\chi}_k\overleftrightarrow{P}^{k,m}_R(\mu'). \end{aligned}$$

Equation (5.118) is valid for the arbitrary form of the phase matrix. But for its widely used block diagonal form (Van de Hulst, 1957; Deirmendjian, 1969; Hovenier et al., 2004) one can simplify (5.118) by introducing the following matrices (Siewert, 2000)

$$\begin{aligned}
\varphi_1(\varphi) &= \mathrm{diag}\,[\cos\varphi\,,\ \cos\varphi\,,\ \sin\varphi\,,\ \sin\varphi]\,,\\
\varphi_2(\varphi) &= \mathrm{diag}\,[-\sin\varphi\,,\ -\sin\varphi\,,\ \cos\varphi\,,\ \cos\varphi]\,,\\
\overleftrightarrow{D}_1 &= \mathrm{diag}\,[\,1\ \ 1\ \ 0\ \ 0\,]\,,\\
\overleftrightarrow{D}_2 &= \mathrm{diag}\,[\,0\ \ 0\ \ 1\ \ 1\,]\,.
\end{aligned}$$

Using it we simplify the scattering integral as

$$\begin{aligned}
\vec{I}_{SP} \;=\; & \frac{\omega_0}{4\pi}\oint\sum_{k=0}^{\infty}(2k+1)\sum_{m=0}^{k}(2-\delta_{0,m})\times\\
& \times\left[\overleftrightarrow{\varphi}_1(m(\varphi-\varphi'))\overleftrightarrow{A}{}^{m}_{k}(\mu,\mu')\overleftrightarrow{D}_1+\overleftrightarrow{\varphi}_2(m(\varphi-\varphi'))\overleftrightarrow{A}{}^{m}_{k}(\mu,\mu')\overleftrightarrow{D}_2\right],
\end{aligned}\tag{5.119}$$

where $\overleftrightarrow{A}{}^{k}_{m}(\mu,\mu') = \overleftrightarrow{\Pi}{}^{k}_{m}(\mu)\overleftrightarrow{\chi}_k(\tau)\overleftrightarrow{\Pi}{}^{k}_{m}(\mu')$, and real-numbered matrix polynomials have the form

$$\overleftrightarrow{\Pi}{}^{k}_{m}(\mu)=\begin{bmatrix} Q^m_k(\mu) & 0 & 0 & 0\\ 0 & R^m_k(\mu) & -T^m_k(\mu) & 0\\ 0 & -T^m_k(\mu) & R^m_k(\mu) & 0\\ 0 & 0 & 0 & Q^m_k(\mu)\end{bmatrix}. \tag{5.120}$$

Following (5.119) we write the form of the solution as

$$\tilde{L}(\tau,\hat{\boldsymbol{l}})=\sum_{m=0}^{\infty}(2-\delta_{0,m})\left[\varphi_1(m\varphi)\vec{L}_1(\tau,\mu)+\varphi_2(m\varphi)\vec{L}_2(\tau,\mu)\right], \tag{5.121}$$

But before solving (5.113) we must bring it and the MSH into the same frame of reference in order to formulate the source function and the bottom boundary condition upon the MSH. The normal frame of reference is very convenient. So we must transform the MSH (5.110) to the normal frame of reference. In order to transform the MSH from the 'along incident ray' frame of reference $\{\nu,\,\psi\}$ to the normal frame of reference $\{\mu,\,\varphi\}$ we use the rotator. We omit some transformations and write:

$$\vec{L}_{MSH}(\tau,\hat{\boldsymbol{l}})=\sum_{c=1,2}\sum_{k=0}^{\infty}\frac{2k+1}{4\pi}\sum_{m=0}^{k}(2-\delta_{0,m})\overleftrightarrow{\varphi}_c(m\varphi)\overleftrightarrow{B}{}^{k}_{m}(\mu,\mu_0,\tau)\overleftrightarrow{D}_c\vec{L}_0\,, \tag{5.122}$$

where $\delta_{0,m}$ is the Kronecker symbol; the exponential matrix $\overleftrightarrow{Z}_k(\tau)$ is defined in (5.109); and $\overleftrightarrow{B}{}^{k}_{m}(\mu,\mu_0,\tau)=\overleftrightarrow{\Pi}{}^{k}_{m}(\mu)\overleftrightarrow{Z}_k(\tau)\overleftrightarrow{\Pi}{}^{k}_{m}(\mu_0)$.

The area of validity of (5.120) is restricted to the block-diagonal form of the phase matrix. Meanwhile the more general form of the phase matrix needs to use

another real-matrix SP-polynomial. We substitute (5.122) into (5.114), differentiate with respect to optical depth, and use the recurrence formula for matrix polynomials $\overleftrightarrow{\Pi}^{k}_{m}(\mu)$ that can be obtained from the recurrence formula for $P^{k}_{m,n}$ and has the form:

$$\mu\overleftrightarrow{\Pi}^{k}_{m}(\mu) = \frac{1}{2k+1}\left[\overleftrightarrow{A}^{m}_{k+1}\overleftrightarrow{\Pi}^{k+1}_{m}(\mu) + \overleftrightarrow{B}^{m}_{k}\overleftrightarrow{\Pi}^{m}_{k}(\mu) + \overleftrightarrow{A}^{m}_{k}\overleftrightarrow{\Pi}^{k-1}_{m}(\mu)\right],$$

$$\left[\overleftrightarrow{A}^{k}_{m}\right]_{rs} = \frac{1}{k}\sqrt{(k^2-m^2)\,(k^2-s^2)}\delta_{rs};$$

$$\left[\overleftrightarrow{B}^{k}_{m}\right]_{rs} = -\frac{2m}{k\,(k+1)}\,(2k+1)\,\delta_{r,-s};\;\; r,s = +0,+2,-2,-0\,.$$

For azimuth recurrence one can easily derive

$$\left[\begin{array}{c} R^{m+1}_{k}(\mu) \\ T^{m+1}_{k}(\mu) \end{array}\right] = \sqrt{\frac{(2m+1)(m+1)}{(2m-2)(m+3)}}\sqrt{1-\mu^2}\left[\begin{array}{c} R^{m}_{k}(\mu) \\ T^{m}_{k}(\mu) \end{array}\right],$$

$$R^{-m}_{k}(\mu) = (-1)^m R^{m}_{k}(\mu);\; T^{-m}_{k}(\mu) = -(-1)^m T^{m}_{k}(\mu)\,.$$

The expressions obtained allow us to use the same route as we did in the scattering integral: we eliminate the complex functions by combining real and imaginary parts and obtain the following form for the source function in the energetic SP-basis and the normal frame of the reference

$$\begin{aligned}\vec{\Delta}(\tau,\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}_0) = {}& \frac{1}{4\pi}\sum_{c=1,2}\sum_{k=0}^{\infty}\sum_{m=0}^{k}(2-\delta_{0,m})\times \\ \times\Bigg[{}& \frac{1}{\mu_0}\overleftrightarrow{\varphi}_c(m\varphi)\left\{\overleftrightarrow{A}^{m}_{k+1}\overleftrightarrow{\Pi}^{k+1}_{m}(\mu)+\overleftrightarrow{B}^{m}_{k}\overleftrightarrow{\Pi}^{m}_{k}(\mu)+\overleftrightarrow{A}^{m}_{k}\overleftrightarrow{\Pi}^{k-1}_{m}(\mu)\right\}(\overleftrightarrow{1}-\Lambda\overleftrightarrow{\chi}_k)\overleftrightarrow{Z}_k\overleftrightarrow{\Pi}^{k}_{m}(\mu_0)\overleftrightarrow{D}_c- \\ & -\overleftrightarrow{\varphi}_c(m\varphi)(2k+1)\overleftrightarrow{\Pi}^{k}_{m}(\mu)(\overleftrightarrow{1}-\overleftrightarrow{\chi}_k)\overleftrightarrow{Z}_k\overleftrightarrow{\Pi}^{k}_{m}(\omega_0\mu_0)\overleftrightarrow{D}_c\Bigg]\vec{L}_0\,. \end{aligned} \tag{5.123}$$

The source function (5.123), together with the previously evaluated scattering integral (5.113), give us the system of equations for DOM expanded in hemispheres, that can be written as usual for each of m-azimuth terms independently:

$$\begin{aligned}\mu^{\pm}_{i}\frac{\partial}{\partial\tau}\vec{L}^{m}_{c}(\tau,\mu^{\pm}_{i}) = {}& -\vec{L}^{m}_{c}(\tau,\mu^{\pm}_{i}) + \frac{\omega_0}{4}\sum_{k=0}^{K}\sum_{j=1}^{N/2}(2k+1)w_j\times \\ & \times\left[\overleftrightarrow{\Pi}^{k}_{m}(\mu^{\pm}_{i})\overleftrightarrow{\chi}_k\overleftrightarrow{\Pi}^{k}_{m}(\mu^{-}_{j})\vec{L}^{m}_{c}(\tau,\mu^{-}_{j}) + \overleftrightarrow{\Pi}^{k}_{m}(\mu^{\pm}_{i})\overleftrightarrow{\chi}_k\overleftrightarrow{\Pi}^{k}_{m}(\mu^{+}_{j})\vec{L}^{m}_{c}(\tau,\mu^{+}_{j})\right] + \\ & + \vec{\Delta}^{m}_{c}(\tau,\mu^{\pm}_{i})\,. \end{aligned} \tag{5.124}$$

The form of (5.124) gives an opportunity to use a matrix operator method completely. We will not repeat it here – the reader can be referred to the scalar case. But we only mention here that the solution for the Milne–Ambartsumian problem

for the arbitrary scattering matrix can be obtained by the matrix-operator method in the following form (with the subsequent azimuth-index summation)

$$\vec{L}_{-}^{m}(0) = -\overleftrightarrow{u}_{12}^{-1}\vec{S}_{-}^{m}\,. \tag{5.125}$$

Now we come to some numerical examples. Fig. 5.4 shows the zenith angle θ dependence for the Q-component of the Stokes vector for the radiance transmitted through the slab. For the sake of the numerical modeling it is convenient to use the special form of the phase matrix (Zege and Chaikovskaya, 1981)

$$\overleftrightarrow{p}_{HG}(\gamma) = p_{HG}(\gamma)\begin{bmatrix} 1 & P_m(1-\gamma^2) & 0 & 0 \\ P_m(1-\gamma^2) & 0.25\,(1+\gamma)^2 & 0 & 0 \\ 0 & 0 & 0.25\,(1+\gamma)^2 & Q_m(1-\gamma^2) \\ 0 & 0 & -Q_m(1-\gamma^2) & 1 \end{bmatrix}, \tag{5.126}$$

that is called the Henyey–Greenstein scattering matrix.

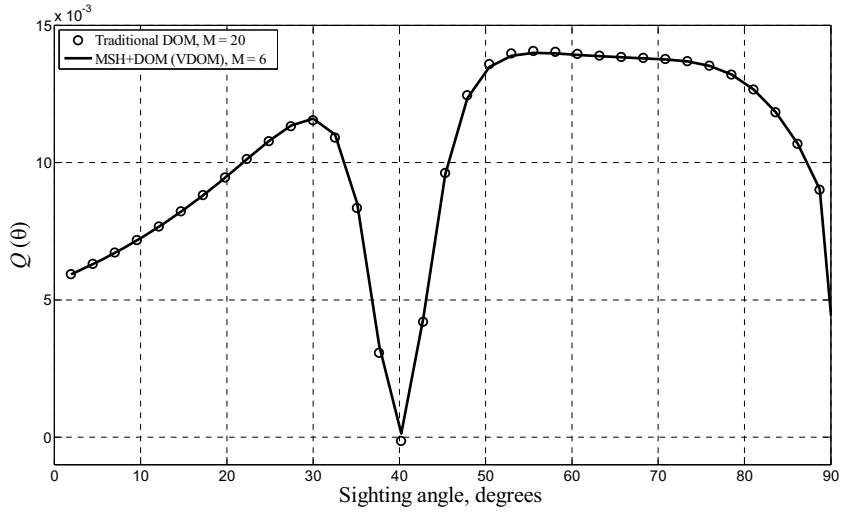


Fig. 5.4. The Q-component is evaluated by the classical and new approaches.

The Henyey–Greenstein scattering matrix is unrealistic, but it is very useful in the numerical simulation of the polarization field in scattering media, because it depends on only three parameters: g, P_m, Q_m are the average scattering cosine, the maximum linear and circular polarization degree within a single scattering act respectively.

In our calculation the Henyey–Greenstein scattering matrix was used with the parameters: $g = 0.9$, $P_m = 0.5$, $Q_m = 0$. The total optical thickness $\tau = 1$ and the single scattering albedo is $\Lambda = 0.8$. The slab is illuminated by natural light and the angle of incidence is $\theta_0 = 40°$. The solid line represents the proposed method MSH + DOM and circle-markers give the classical DOM solution with the only singular part subtracted. As one can see from Fig. 5.4, the proposed method is not

so sensitive to azimuth m-expansion: $M = 20$ for azimuth harmonics are needed to obtain the classical solution and only $M = 6$ for our method with anisotropic part subtraction. The minimum azimuthal harmonics was obtained from the fact that for this optical thickness the Q-component for exactly forward scattered light must be zero. The bottom boundary is assumed to be non-reflecting.

The results for polarization degree defined with respect to parallel (l) and perpendicular (r) components as

$$p(\theta) = \frac{L_l(\theta) - L_r(\theta)}{L_l(\theta) + L_r(\theta)} = \frac{Q(\theta)}{I(\theta)} \tag{5.127}$$

are shown in Fig. 5.5 (backward hemisphere of the top boundary – the radiation reflected from the slab). The reflectance coefficients are: $\rho = 0$ (solid line), $\rho = 0.3$ (dash-dot), $\rho = 0.7$ (dashed), and $\rho = 1$ (dots). One can see that depolarization properties of the bottom reflectance greatly influence the polarization state of radiation. In this case the Haze L model is used for the approximation of the phase function, $\omega_0 = 0.9$ and $\tau = 1$.

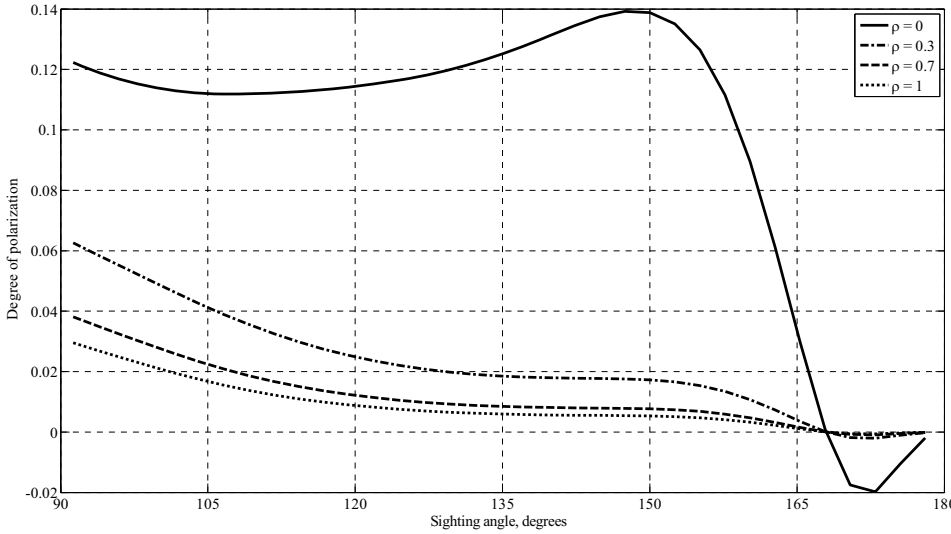


Fig. 5.5. Degree of polarization vs. phase angle within the backward hemisphere for Mie scattering – Haze L with different reflectance coefficients of the underlying surface ρ. The phase angle is the angle between the direction of slab irradiance and the direction of observation.

In Fig. 5.6 Rayleigh scattering for different single scattering albedos of the slab is shown. We set $\tau = 1$, $\omega_0 = 0.9$ (dash-dot), $\omega_0 = 0.5$ (dashed), $\omega_0 = 0.1$ (solid). Angle φ is the phase angle between the direction of irradiance of the slab and the direction of observation. The lines in the figure are polar-phase curves. The charts demonstrate the action of Umov (1905): the higher the surface albedo, the lower the degree of polarization.

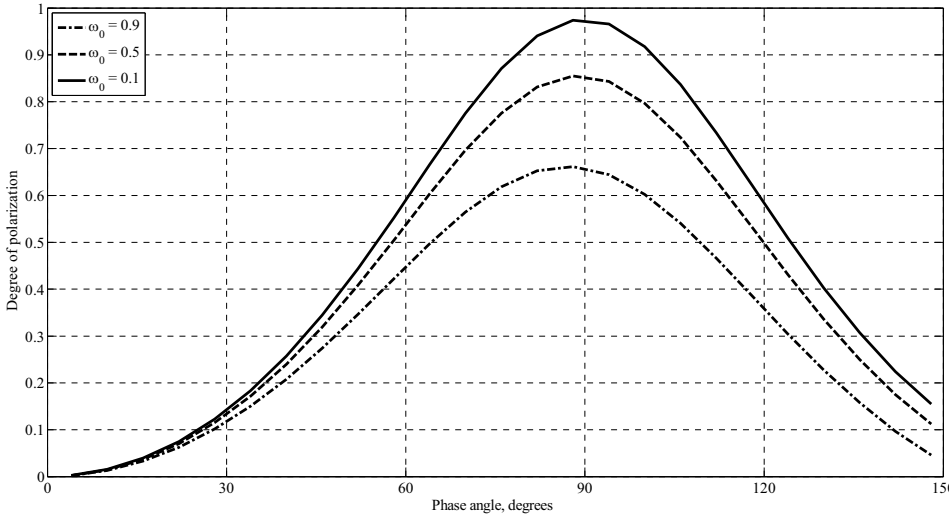


Fig. 5.6. Degree of polarization, Rayleigh scattering, $\tau = 1$, natural light illumination, polar curves in dependence on the different single scattering albedos.

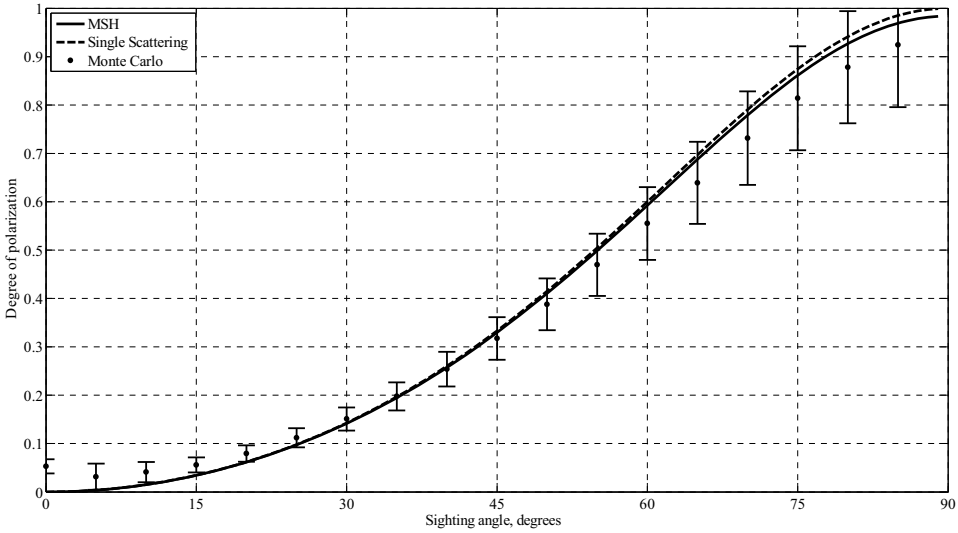


Fig. 5.7. Degree of polarization dependence for Rayleigh scattering: MSH computational aspects.

And finally Fig. 5.7 shows the zenith distribution of the degree of polarization at the optical depth $\tau = 0.1$ for quite smooth Rayleigh scattering matrix obtained by the MSH itself. In spite of smooth scattering there is a good agreement because of the boundary problem singularity. The MSH describes some known facts for Rayleigh scattering: the greater the scattering angle (to be exact the zenith angle of observation) the higher the polarization state of the scattered radiation. The radiation scattered forward for the normal irradiance is nonpolarized and the de-

gree of polarization culmination for the Rayleigh scattering is 90°. An insignificant deviation of the MSH and the single scattering occurred together with the sight angle increasing caused by the increase of the optical thickness for the inclined directions (the greater the inclination the higher the deviation).

5.9 MSH in arbitrary medium geometry

In order to generalize the above-mentioned approach upon the arbitrary three-dimensional (3-D) geometry of the scattering medium it is necessary to find the anisotropic part solution for a point unidirectional source of light radiation. A point unidirectional source provides the most general fundamental form for the RTE boundary problem solution with arbitrary boundary conditions. In order to obtain the anisotropic part it is enough to explore the radiation distribution from the point unidirectional source in an infinite medium since the anisotropic part is determined by the local value of the medium situated close to the line of sight of the source from a detector position.

Let us take the following frame of reference in the phase space of the radiative transfer (Fig. 5.8):

$$\eta = (\hat{\boldsymbol{q}}, \hat{\boldsymbol{r}}) = \cos\alpha; \quad \mu = (\hat{\boldsymbol{l}}, \hat{\boldsymbol{r}}) = \cos\beta; \; r; \; \varphi$$

where $\hat{\boldsymbol{q}}$ is the emission direction of the point unidirectional source.

Accordingly, $L(\boldsymbol{r}, \hat{\boldsymbol{l}}) = L(r, \eta, \mu, \varphi)$ and we could write for the differential operator of RTE (5.1):

$$(\hat{\boldsymbol{l}}, \nabla)L = \frac{dL}{ds} = \frac{\partial L}{\partial r}\frac{dr}{ds} + \frac{\partial L}{\partial \alpha}\frac{d\alpha}{ds} + \frac{\partial L}{\partial \beta}\frac{d\beta}{ds} + \frac{\partial L}{\partial \varphi}\frac{d\varphi}{ds},$$

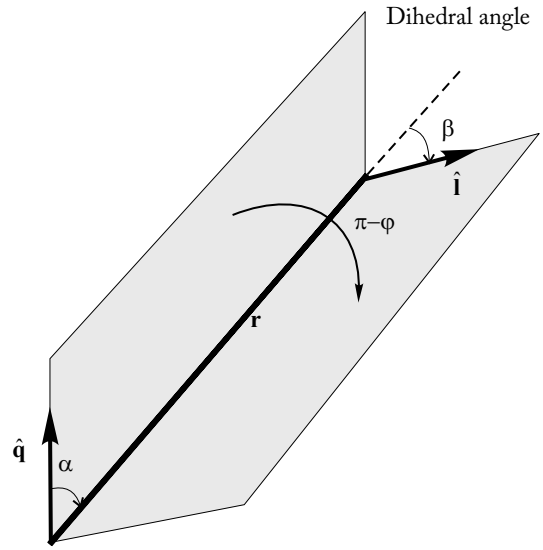


Fig. 5.8. The geometry of the boundary value problem for the point unidirectional source.

where s is the distance along the ray (the natural frame of reference).

It is not too hard to see that

$$\frac{dr}{ds} = (\hat{\boldsymbol{l}}, \nabla) r = (\hat{\boldsymbol{l}}, \hat{\boldsymbol{r}}) = \mu\,;$$

$$\begin{aligned}\frac{d\alpha}{ds} &= \frac{1}{\sin\alpha}\frac{d\cos\alpha}{ds} = -\frac{1}{\sin\alpha}(\hat{\boldsymbol{l}}, \nabla)\frac{(\hat{\boldsymbol{q}}, \boldsymbol{r})}{r} = -\frac{1}{\sin\alpha}\left(\hat{\boldsymbol{l}}, -\frac{1}{r^2}(\hat{\boldsymbol{l}}, \hat{\boldsymbol{r}}) + \frac{\hat{\boldsymbol{q}}}{r}\right) = \\ &= -\frac{1}{\sqrt{1-\eta^2}}\left(-\frac{\eta\mu}{r} + \frac{\eta\mu - \sqrt{(1-\eta^2)(1-\mu^2)}\cos\varphi}{r}\right) = \frac{\sqrt{(1-\mu^2)}}{r}\cos\varphi\,;\end{aligned}$$

$$\begin{aligned}\frac{d\beta}{ds} &= -\frac{1}{\sin\beta}\frac{d\cos\beta}{ds} = -\frac{1}{\sqrt{1-\mu^2}}(\hat{\boldsymbol{l}}, \nabla)\frac{(\hat{\boldsymbol{l}}, \boldsymbol{r})}{r} = -\frac{1}{\sin\beta}\left(\hat{\boldsymbol{l}}, -\frac{1}{r^2}(\hat{\boldsymbol{l}}, \hat{\boldsymbol{r}}) + \frac{1}{r}\right) \\ &= -\frac{\sqrt{1-\mu^2}}{r}\,;\end{aligned}$$

$$\begin{aligned}\frac{d\varphi}{ds} &= \frac{1}{\sin\alpha}\frac{d}{ds}\frac{\cos\gamma - \cos\alpha\cos\beta}{\sin\alpha\sin\beta} = \\ &= \frac{\frac{d\alpha}{ds}(\sin\alpha\,\cos\beta + \cos\alpha\,\sin\beta\,\cos\varphi) - \frac{d\beta}{ds}(\cos\alpha\,\sin\beta - \sin\alpha\,\cos\beta\,\cos\varphi)}{\sin\varphi\,\sin\alpha\,\sin\beta} \\ &= -\frac{\cos\alpha\,\sin\beta}{r\,\sin\alpha}\sin\varphi = -\frac{\sqrt{1-\mu^2}}{\sqrt{1-\eta^2}}\frac{\sin\varphi}{r}\eta\,;\end{aligned}$$

where $\cos\gamma = (\hat{\boldsymbol{q}}, \hat{\boldsymbol{l}}) = \mu\eta - \sqrt{(1-\eta^2)(1-\mu^2)}\cos\varphi$. We also take into account that $\cos\gamma$ doesn not change along the ray.

Based on the obtained results we can write the RTE for the point unidirectional source:

$$\begin{aligned}&\mu\frac{\partial L}{\partial r} + \frac{1-\mu^2}{r}\frac{\partial L}{\partial \mu} - \frac{\sqrt{1-\mu^2}}{r}\left(\sqrt{1-\eta^2}\frac{\partial L}{\partial \eta}\,\cos\varphi + \frac{\eta}{\sqrt{1-\eta^2}}\frac{\partial L}{\partial \varphi}\,\sin\varphi\right) \\ &\qquad = -k_{ext}L(r,\eta,\mu,\varphi) + \frac{\omega_0 k_{ext}}{4\pi}\oint L(r,\eta,\mu',\varphi')p(\hat{\boldsymbol{l}}, \hat{\boldsymbol{l}}')d\hat{\boldsymbol{l}}'\,. \qquad (5.128)\end{aligned}$$

Let us represent the required solution as the expansion in series on the spherical functions

$$\begin{aligned}&L(r,\eta,\mu,\varphi) \\ &= \frac{1}{2\pi}\sum_{l=0}^{\infty}\sum_{k=0}^{\infty}\sum_{m=-\infty}^{\infty}\frac{2k+1}{2}\frac{2l+1}{2}\sqrt{\frac{(k-m)!}{(k+m)!}}\sqrt{\frac{(l-m)!}{(l+m)!}}C_{kl}^{m}(r)P_k^m(\mu)P_l^m(\eta)\mathrm{e}^{im\varphi}\,, \qquad (5.129)\end{aligned}$$

or in the equivalent form

$$L(r,\eta,\mu,\varphi)$$
$$= \frac{1}{\sqrt{2\pi}} \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \sum_{m=-\infty}^{\infty} \sqrt{\frac{2k+1}{2}} \frac{2l+1}{2} \sqrt{\frac{(l-m)!}{(l+m)!}} C_{kl}^{m}(r) Y_{km}(\hat{\boldsymbol{l}}) P_l^m(\eta)\, \mathrm{e}^{im\varphi} \tag{5.130}$$

$$= \frac{1}{\sqrt{2\pi}} \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \sum_{m=-\infty}^{\infty} \sqrt{\frac{2l+1}{2}} \frac{2k+1}{2} \sqrt{\frac{(k-m)!}{(k+m)!}} C_{kl}^{m}(r) Y_{lm}(\hat{\boldsymbol{q}}) P_k^m(\mu) \mathrm{e}^{im\varphi}. \tag{5.131}$$

We substitute the radiance distribution (5.129) and the phase function Legendre polynomials expansion in the RTE (5.128). Then we multiply the RTE (5.128) by $Q_k^m(\mu) Q_l^m(\eta)\, \mathrm{e}^{-im\varphi}$ and integrate it over the complete space for μ, η, φ. After that the orthogonal properties for the spherical functions are employed to obtain the infinite system of the coupled differential equations for the spherical harmonics method

$$\begin{aligned}
&\frac{d}{dr}\left[\sqrt{(k-m)(k+m)}C_{k-1,l}^{m}(r) + \sqrt{(k+m+1)(k-m+1)}C_{k+1,l}^{m}(r)\right] \\
&+\frac{1}{r}\left[(k+2)\sqrt{(k+m+1)(k-m+1)}C_{k+1,l}^{m}(r) - (k-1)\sqrt{(k+m)(k-m)}C_{k-1,l}^{m}(r)\right] \\
&+\frac{1}{2r}\Big\{\sqrt{(l+m)(l-m+1)} \\
&\quad\left[\sqrt{(k-m+1)(k-m+2)}C_{k+1,l}^{m-1}(r) - \sqrt{(k+m)(k+m-1)}C_{k-1,l}^{m-1}(r)\right] \\
&\quad+\left[\sqrt{(k+m+1)(k+m+2)}C_{k+1,l}^{m+1}(r) - \sqrt{(k-m-1)(k-m)}C_{k-1,l}^{m+1}(r)\right] \\
&\sqrt{(l-m)(l+m+1)}\Big\} = -(2k+1)k_{ext}(1-\omega_0 p_k)C_{kl}^{m}(r)\,.
\end{aligned} \tag{5.132}$$

Let us use the main features of the MSH method to solve the obtained system for the point unidirectional source (5.132). In such a case, the generalization of MSH is not trivial since the system has two expansion indices: k, l. That is why the subsequent development of MSH method is necessary here. To do this we assume (Budak et al., 2008) the following:

– the continuous dependence of the expansion coefficients in (5.128) on the harmonics number, i.e.
$$C_{kl}^{m}(r) = C^{m}(k,l,r)\,;$$
– the weak and monotonic dependence of the function $C^m(k,l,r)$ on k, l, that allows us to present the solution in the following form
$$C^{m}(k\pm1,l,r) \approx C^{m}(k,l,r) \pm \frac{\partial C^{m}(k,l,r)}{\partial k}\,;$$
– the basic contribution in the expansion (5.129) belongs to the terms with $k,l \gg 1$ and $k,l \gg m$, so the following transformations are justified:
$$\sqrt{(l+m)(l-m+1)} = \sqrt{l(l+1)-m(m-1)} \approx \sqrt{l(l+1)}\,;$$

$$\sqrt{(l-m)(l+m+1)} = \sqrt{l(l+1)-m(m+1)} \approx \sqrt{l(l+1)}\,;$$
$$\sqrt{(k+m)(k-m)} = \sqrt{k^2-m^2} \approx k\,;$$

$$\begin{aligned}\sqrt{(k-m+1)(k+m+1)} &= \sqrt{k(k+1)+(m-1)^2+(m-1)(2k+1)} \\ &\approx \sqrt{k(k+1)}+(m-1)\frac{2k+1}{2\sqrt{k(k+1)}}\,;\end{aligned}$$

$$\begin{aligned}\sqrt{(k-m+1)(k-m+2)} &= \sqrt{k(k+1)+(m-1)^2-(m-1)(2k+1)} \\ &\approx \sqrt{k(k+1)}-(m-1)\frac{2k+1}{2\sqrt{k(k+1)}}\,;\end{aligned}$$

$$\begin{aligned}\sqrt{(k-m-1)(k-m)} &= \sqrt{k(k+1)+(m-1)^2-(m+1)(2k+1)} \\ &\approx \sqrt{k(k+1)}-(m+1)\frac{2k+1}{2\sqrt{k(k+1)}}\,;\end{aligned}$$

$$\begin{aligned}\sqrt{(k+m+1)(k-m+2)} &= \sqrt{k(k+1)+(m-1)^2+(m+1)(2k+1)} \\ &\approx \sqrt{k(k+1)}+(m+1)\frac{2k+1}{2\sqrt{k(k+1)}}\,.\end{aligned}$$

We use these assumptions in equation (5.132) to represent it as follows

$$\begin{aligned}&\frac{\partial Y^m}{\partial r}+\frac{2k(k+1)}{2k+1}\frac{1}{r}\frac{\partial Y^m}{\partial k}+\frac{\sqrt{l(l+1)}}{2r}\left[\frac{2\sqrt{k(k+1)}}{2k+1}\times\right. \\ &\qquad\left.\times\left(\frac{\partial Y^{m+1}}{\partial k}+\frac{\partial Y^{m-1}}{\partial k}\right)\frac{1}{\sqrt{k(k+1)}}\left((m+1)Y^{m+1}-(m-1)Y^{m-1}\right)\right] \\ &=-k_{ext}(1-\omega_0 p_k)Y^m(k,l,r),\end{aligned} \tag{5.133}$$

with

$$C^m(k,l,r)=\frac{Y^m(k,l,r)}{r^2}\,.$$

Now we change the following variables

$$\lambda=\sqrt{l(l+1)}\,;\kappa=\sqrt{k(k+1)}\,;$$

Then with the representation

$$\frac{\partial}{\partial k}=\frac{2k+1}{2\sqrt{k(k+1)}}\frac{\partial}{\partial \kappa}\,,$$

we get an equation:

$$\begin{aligned}&\frac{\partial Y^m}{\partial r}+\frac{\kappa}{r}\frac{\partial Y^m}{\partial \kappa}+\frac{\lambda}{2r}\left[\frac{\partial Y^{m+1}}{\partial \kappa}+\frac{\partial Y^{m-1}}{\partial \kappa}+\frac{1}{\kappa}\left((m+1)Y^{m+1}-(m-1)Y^{m-1}\right)\right] \\ &\quad=-k_{ext}(1-\omega_0 p(k))Y^m(k,l,r)\,.\end{aligned} \tag{5.134}$$

We introduce in our evaluations a new function depending on the two vectors in a plain $f(r, \boldsymbol{\kappa}, \boldsymbol{\lambda})$; the solution $Y^m(r,\kappa,\lambda)$ is the Fourier ψ-spectrum between the vectors $\boldsymbol{\kappa}$ and $\boldsymbol{\lambda}$, i.e.

$$f(r,\kappa,\lambda) = \sum_{m=-\infty}^{m=+\infty} Y^m(r,\kappa,\lambda)\,\mathrm{e}^{im\psi}\,. \tag{5.135}$$

One can easily notice the following properties for the expansion series (5.135):

$$\begin{aligned}
&\sum_{m=-\infty}^{m=+\infty} Y^{m\pm1}(r,\kappa,\lambda)\mathrm{e}^{im\psi} = \mathrm{e}^{\mp i\psi} f(r,\kappa,\lambda),\\
&\sum_{m=-\infty}^{m=+\infty} (m\pm)Y^{m\pm1}(r,\kappa,\lambda)\,\mathrm{e}^{im\psi} = -i\,\mathrm{e}^{\mp i\psi}\frac{\partial f(r,\kappa,\lambda)}{\partial\psi}\,.
\end{aligned} \tag{5.136}$$

Using the properties of the Fourier expansion we can obtain the expression for $Y^m(r,\kappa,\lambda)$ via $f(r,\boldsymbol{\kappa},\boldsymbol{\lambda})$:

$$Y^m(r,\kappa,\lambda) = \frac{1}{2\pi}\int_0^{2\pi} f(r,\kappa,\lambda)\,\mathrm{e}^{-im\psi}\,d\psi\,.$$

We multiply equation (5.134) by $\mathrm{e}^{im\psi}$ and sum it up with m from $-\infty$ to $+\infty$. Then with the properties (5.136) we have the relation

$$\begin{aligned}
&\frac{\lambda}{2r}\sum_{m=-\infty}^{m=+\infty}\left[\frac{\partial Y^{m+1}}{\partial\kappa}+\frac{\partial Y^{m-1}}{\partial\kappa}+\frac{1}{\kappa}\left((m+1)Y^{m+1}-(m-1)Y^{m-1}\right)\right]\mathrm{e}^{im\psi} =\\
&\quad=\frac{\lambda}{2r}\left[\mathrm{e}^{i\psi}\frac{\partial f}{\partial\kappa}+\mathrm{e}^{-i\psi}\frac{\partial f}{\partial\kappa}+\frac{1}{\kappa}\left(i\,\mathrm{e}^{i\psi}\frac{\partial f}{\partial\psi}-i\,\mathrm{e}^{-i\psi}\frac{\partial f}{\partial\psi}\right)\right]\\
&\quad=\frac{\lambda}{r}\left(\cos\psi\frac{\partial f}{\partial\kappa}-\frac{\sin\psi}{\kappa}\frac{\partial f}{\partial\psi}\right)=\frac{1}{r}(\lambda,\nabla)f(r,\kappa,\lambda)\,.
\end{aligned}$$

that allows us to transform equation (5.134) to the following form

$$\frac{\partial f}{\partial r}+\frac{\kappa}{r}\frac{\partial f}{\partial\kappa}+\frac{1}{r}(\lambda,\nabla)f=-k_{ext}(1-\omega_0 p(\kappa))f(r,\kappa,\lambda)\,. \tag{5.137}$$

Note that the relation that appears to be obvious

$$\frac{\kappa}{r}\frac{\partial f}{\partial\kappa}=\frac{1}{r}(\kappa,\nabla)f(r,\kappa,\lambda)\,,$$

provides the transformation of (5.137) to the following form

$$\frac{\partial f}{\partial r}+\frac{1}{r}(\kappa+\lambda,\nabla)f=-k_{ext}(1-\omega_0 p(\kappa))f(r,\kappa,\lambda)\,. \tag{5.138}$$

We seek the solution in the following form

$$f(r,\kappa,\lambda)=\exp\left(-k_{ext}r+k_{ext}\omega_0 rF(\kappa,\lambda)\right)\,. \tag{5.139}$$

From (5.137) we have an equation

$$F + (\kappa + \lambda, \nabla)F(\kappa, \lambda) = p(\kappa)\,. \tag{5.140}$$

We will solve equation (5.140) by the integration along the characteristic curve

$$\kappa' = -\lambda + \frac{\kappa + \lambda}{|\kappa + \lambda|}\xi\,,$$

and transform (5.140) to an ordinary differential equation

$$F(\xi) + \xi\frac{dF}{d\xi} = p\left(\left|-\lambda + \frac{\kappa + \lambda}{|\kappa + \lambda|}\xi\right|\right)\,,$$

with the solution in the following form

$$F(\kappa, \lambda) = \frac{1}{|\kappa + \lambda|}\int\limits_0^{|\kappa+\lambda|} p\left(\left|-\lambda + \frac{\kappa + \lambda}{|\kappa + \lambda|}\xi\right|\right)\,d\xi + \frac{C_0}{|\kappa + \lambda|},$$

The continuity condition of $F(\kappa,\lambda)$ for $\boldsymbol{\kappa} = -\boldsymbol{\lambda}$ gives us $C_0 = 0$.

From this result and (5.139) we get

$$f(r, \kappa, \lambda) = \exp\left[-k_{ext}r + k_{ext}\omega_0 r\int\limits_0^1 p\left(|(\kappa + \lambda)\zeta - \lambda|\right)\,d\zeta\right]\,.$$

From this equation we can write:

$$C_{kl}^m(r) = \frac{\mathrm{e}^{-k_{ext}r}}{2\pi r^2}\int\limits_0^{2\pi} \exp\left(-im\psi + \omega_0 k_{ext} r\chi(k, l, \psi)\right)\,d\psi, \tag{5.141}$$

with

$$\chi(k,l,\psi) \equiv \int\limits_0^1 p\left(\sqrt{l(l+1)(1-\zeta)^2 + k(k+1)\zeta^2 - 2\zeta(1-\zeta)\sqrt{l(l+1)k(k+1)}\,\cos\psi}\right)\,d\zeta.$$

The equation obtained is quite difficult for direct analysis. We evaluate the radiation distribution for a point isotropic source:

$$L(r, \mu) = \sum_{k=0}^{\infty}\frac{2k + 1}{4\pi}C_k(r)P_k(\mu)\,, \tag{5.142}$$

where $L(\boldsymbol{r}, \hat{\boldsymbol{l}}) = L(r, \mu)$, $\mu = (\hat{\boldsymbol{l}}, \boldsymbol{r})\Big/ r$.

The coefficients $C_k(r)$ for the point isotropic source (5.142) have the following evident relation with the coefficients $C_{kl}^m(r)$ for point unidirectional:

$$C_k(r) = C_{k0}^0(r) = \frac{\mathrm{e}^{-k_{ext}r}}{r^2}\exp\left[\frac{\omega_0 k_{ext} r}{\sqrt{k(k+1)}}\int\limits_0^{\sqrt{k(k+1)}} p(\zeta)\,d\zeta\right]\,. \tag{5.143}$$

In order to make the analytical analysis of the obtained solution features we expand (5.140) with respect to the orders of scattering that is equivalent to the expansion of (5.143) in a Taylor series for the powers of $\omega_{\mathrm{w}} k_{ext} r$. For the first three orders of scattering we have

$$L_0(r,\mu) = \sum_{k=0}^{\infty} \frac{2k+1}{4\pi r^2}\, \mathrm{e}^{-k_{ext} r} P_k(\mu) = \frac{\mathrm{e}^{-k_{ext} r}}{2\pi r^2} \delta(1-\mu)\,, \tag{5.144}$$

$$L_1(r,\mu) = \sum_{k=0}^{\infty} \frac{2k+1}{4\pi r^2}\, \mathrm{e}^{-\varepsilon r} \left\{ \frac{k_{sca} r}{q} \int_0^q p(\xi)\, d\xi \right\} P_k(\mu),\ q = \sqrt{k(k+1)}\,, \tag{5.145}$$

$$L_2(r,\mu) = \sum_{k=0}^{\infty} \frac{2k+1}{4\pi r^2}\, \mathrm{e}^{-\varepsilon r} \left\{ \frac{k_{sca} r}{q} \int_0^q p(\xi)\, d\xi \right\}^2 P_k(\mu)\,. \tag{5.146}$$

As can be seen from (5.144) the zero-order scattering has the δ-singularity of the RTE solution. As far as we are interested in the 'behavior' of (5.145) and (5.146) in the area around the singular point $\mu = 1$, the basic contribution could be given by the terms with $k \gg 1$. According to the Debye localization principle in Mie theory (Deirmendjian, 1969) one can write the relation for the case of an arbitrary phase function

$$\int_0^q x(\zeta)\, d\zeta \underset{k\to\infty}{\longrightarrow} C_0\,,$$

where C_0 is a constant value.

So, for the case $\mu \sim 1$ the expansion series (5.145) and (5.146) have the following forms

$$\begin{aligned} L_1(r,\mu) &= \frac{k_{sca} C_0}{2\pi r}\, \mathrm{e}^{-k_{ext} r} \sum_{k=0}^{\infty} P_k(\mu), \\ L_2(r,\mu) &= \frac{k_{sca}^2 C_0^2}{4\pi}\, \mathrm{e}^{-k_{ext} r} \sum_{k=0}^{\infty} \left(\frac{1}{k+1} + \frac{1}{k} \right) P_k(\mu)\,, \end{aligned}$$

with the approximation $2\sqrt{k(k+1)} \approx 2k+1$ at $k \gg 1$.

It is not difficult to sum up the above-mentioned series using the properties of the generic function for the Legendre polynomials. As a result we have

$$L_1(r,\mu) = \frac{k_{sca} C_0}{2\pi r} \mathrm{e}^{-k_{ext} r} \frac{1}{\sqrt{2(1-\mu)}}, L_2(r,\mu) = \frac{k_{sca}^2 C_0^2}{2\pi r}\, \mathrm{e}^{-k_{ext} r} \left(-\ln(1-\mu)\right)\,.$$

Comparing the result obtained with the one from the Table 5.1 that shows the singularities of the exact solution (Germogenova, 1986) it is possible to conclude that the RTE solution in MSH (5.143) keeps all the singularities of the exact solution. In the same way, it is not hard to show that the orders of scattering of more than two are the smooth functions. Therefore the MSH describes all the orders of scattering with a different accuracy. Thus, the zero-order scattering (direct non-scattered light) is described exactly and the scattering of higher orders is described approximately.

5.10 Regular part computation in arbitrary medium geometry

The representation of (5.8) indeed makes the rest $\tilde{L}(\boldsymbol{r},\hat{\boldsymbol{l}})$ a smooth, slowly varying function in the space of its arguments. Our analysis has shown that in practically any case the zenith directions amount corresponds to $N < 51$, and the azimuth harmonics amount to $M < 8$. According to a finite element method (Mitchell and Wait, 1976), it allows us to construct in the space the tetrahedron or hexahedron mesh to store in every vertex the discrete ordinate of the decomposition coefficients in the series of the azimuth angular distribution. The values between the nodes are calculated using one of the approximation schemes (Mitchell and Wait, 1976). The suggested method of RTE solution allows us to solve problems in the transport theory in media with the arbitrary 3-D geometry. However, the total number of equations in the gained set can be as large as hundreds of thousands.

The best technique for the solution of such a coupled system of equations is the method of iterations. This could be done by employing an integral transfer equation obtained from the boundary-value problem by the formal solution in the assumption of the known right-hand member.

After the anisotropic part elimination

$$L(\boldsymbol{r},\hat{\boldsymbol{l}}) = L_a(\boldsymbol{r},\hat{\boldsymbol{l}}) + \tilde{L}(\boldsymbol{r},\hat{\boldsymbol{l}}) \tag{5.147}$$

RTE in the arbitrary medium geometry has the form

$$(\hat{\boldsymbol{l}},\nabla)\tilde{L}(\boldsymbol{r},\hat{\boldsymbol{l}}) = -k_{ext}\tilde{L}(\boldsymbol{r},\hat{\boldsymbol{l}}) + \frac{\omega_0 k_{ext}}{4\pi}\oint p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}')\tilde{L}(\boldsymbol{r},\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}' + S(\boldsymbol{r},\hat{\boldsymbol{l}}), \tag{5.148}$$

where

$$S(\boldsymbol{r},\hat{\boldsymbol{l}}) = (\hat{\boldsymbol{l}},\nabla)L_a(\boldsymbol{r},\hat{\boldsymbol{l}}) + k_{ext}L_a(\boldsymbol{r},\hat{\boldsymbol{l}}) - \frac{\omega_0 k_{ext}}{4\pi}\oint p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}')L_a(\boldsymbol{r},\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}'$$

is the anisotropic part misalignment that is easily computed since $L_a(\boldsymbol{r},\hat{\boldsymbol{l}})$ is always presented analytically in MSH as the series on the surface functions.

For the transition to the Peierls integral RTE we deal with the formal solution of (5.148) assuming that the integral term is the known function. We make the transition to the variables along the ray from the point $\mathbf{R}$ in the direction$\hat{\boldsymbol{l}}$. Then an arbitrary point $\mathbf{r}$ on the ray at the distance ξ from $\mathbf{R}$ along the ray must satisfy the relation

$$\boldsymbol{r} = \boldsymbol{R} + \xi\hat{\boldsymbol{l}}\,,$$

As the result, equation (5.148) is:

$$\frac{d}{d\xi}\tilde{L}(\boldsymbol{R}+\xi\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}) + \varepsilon\tilde{L}(\boldsymbol{R}+\xi\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}) = F(\boldsymbol{R}+\xi\hat{\boldsymbol{l}},\hat{\boldsymbol{l}})\,, \tag{5.149}$$

where

$$F(\boldsymbol{r},\hat{\boldsymbol{l}}) \equiv \frac{\omega_0 k_{ext}}{4\pi}\oint p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}')\tilde{L}(\boldsymbol{r},\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}' + S(\boldsymbol{r},\hat{\boldsymbol{l}})\,.$$

The expression (5.149) is the linear inhomogeneous differential equation relative to the variable ξ the solution of which can be written as

$$\tilde{L}(\boldsymbol{R}+\xi\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}) = \Big(L_e(\boldsymbol{R}_0,\hat{\boldsymbol{l}}) - L_a(\boldsymbol{R}_0,\hat{\boldsymbol{l}})\Big)\,\mathrm{e}^{-k_{ext}\xi_0} + \mathrm{e}^{-k_{ext}\xi_0}\int\limits_0^{\xi_0} F(\boldsymbol{R}+\zeta\hat{\boldsymbol{l}},\hat{\boldsymbol{l}})\,\mathrm{e}^{k_{ext}\zeta}\,d\zeta\,, \tag{5.150}$$

where ξ_0 is the distance along the ray from the observation point $\mathbf{r}$ to the medium border; $\mathbf{R}_0$ is the intersection point of the ray with the medium border; $L_e(\boldsymbol{R},\hat{\boldsymbol{l}})$ is the outer radiation incident on the border. The subtraction of $L_a(\boldsymbol{R}_0,\hat{\boldsymbol{l}})$ is associated with the fact that MSH neglects the backscattering, and the radiance in MSH on the border inside the medium is nonzero.

With the known form of the function $F(\boldsymbol{r},\hat{\boldsymbol{l}})$, equation (5.150) can be rewritten in the form of the integral equation

$$\tilde{L}(\boldsymbol{R}+\xi\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}) = L_0(\boldsymbol{R}+\xi\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}) + \frac{\omega_0 k_{ext}}{4\pi}\int\limits_0^{\xi_0}\mathrm{e}^{k_{ext}(\zeta-\xi_0)}\oint p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}')\tilde{L}(\boldsymbol{R}+\zeta\hat{\boldsymbol{l}}',\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}'\,d\zeta\,, \tag{5.151}$$

where

$$L_0(\boldsymbol{R}+\xi\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}) = \Big(L_e(\boldsymbol{R}_0,\hat{\boldsymbol{l}}) - L_a(\boldsymbol{R}_0,\hat{\boldsymbol{l}})\Big)\,\mathrm{e}^{-k_{ext}\xi_0} + \int\limits_0^{\xi_0}\mathrm{e}^{k_{ext}(\zeta-\xi_0)}S(\boldsymbol{R}+\zeta\hat{\boldsymbol{l}},\hat{\boldsymbol{l}})\,d\zeta\,.$$

Equation (5.151) is the Fredholm integral equation of the II kind, the which solution is represented in the Neumann series

$$\tilde{L}(\boldsymbol{r},\hat{\boldsymbol{l}}) = \sum_{n=0}^{\infty} L_n(\boldsymbol{r},\hat{\boldsymbol{l}})\,, \tag{5.152}$$

where

$$L_n(\boldsymbol{R}+\xi\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}) = \frac{\omega_0 k_{ext}}{4\pi}\int\limits_0^{\xi_0}\mathrm{e}^{k_{ext}(\zeta-\xi_0)}\oint p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}')L_{n-1}(\boldsymbol{R}+\zeta\hat{\boldsymbol{l}}',\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}'\,d\zeta\,. \tag{5.153}$$

Every nth term of the series (5.152) represents physically the radiance of the nth order of scattering of the regular part. The expression (5.153) presents the definite integral over the ray by $d\zeta$ of the convolution over the solid angle of RAD of the previous order of scattering with the phase function. To calculate (5.153) let us divide the bulk of medium into the mesh (Fig. 5.9) of the hexagonal cells (Mitchell and Wait, 1976), in each node $\mathbf{r}_{ij}$ of which we will store the radiance discrete values by the fixed direction $\hat{\boldsymbol{l}}_{pq}$ in accordance with the selected quadrature formula of the integration over the solid angle (Koch and Becker, 2004). At first, the convolution over the solid angle is calculated in every node, and when making the calculation of the integral over the ray the integral convolution values are interpolated between the nodes using one of the approximation schemes.

We show in Fig. 5.10 the comparison of the calculation results of the integral convolution over the complete solid angle in the cases of the subtraction from the complete solution of the direct nonscattering radiation only and of the whole the anisotropic part. It is not difficult to see that the proposed approach of the

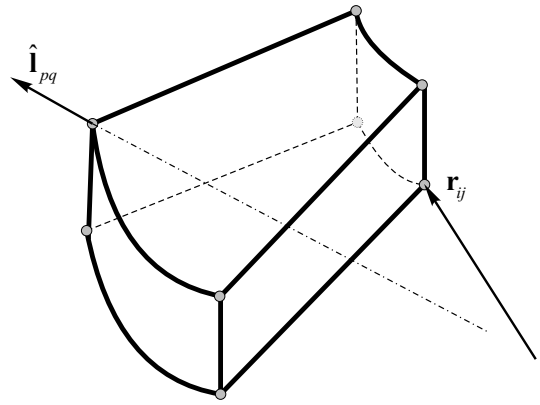


Fig. 5.9. The mesh cell to approximate the radiance values between nodes.

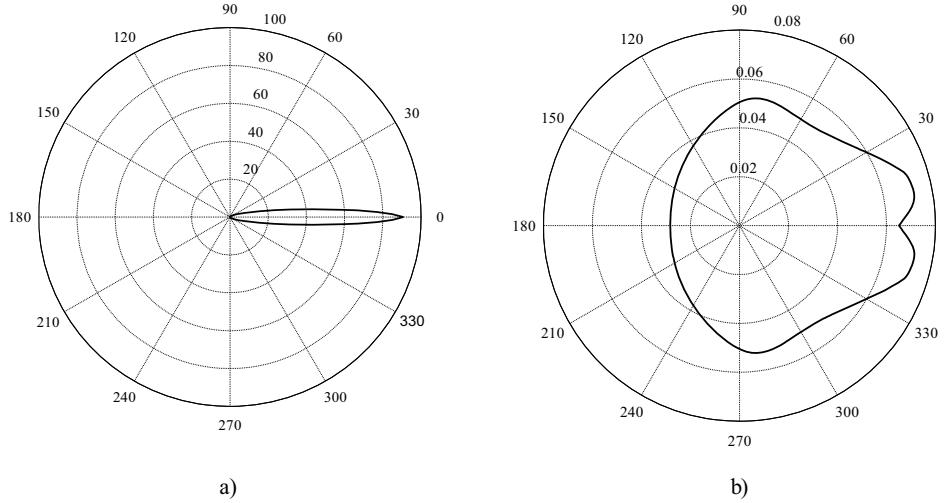


Fig. 5.10. The angular distribution of the scattering integral at the depth $\tau = 1.0$: by eliminating from RAD (a) only the direct radiation, (b) the whole anisotropic part by MSH. The medium parameters: $g = 0.9$, $\omega_0 = 0.8$, $\tau_0 = 5.0$.

whole anisotropic part elimination solution changes the residual part solution into a considerably smoother function. This essentially reduces requirements for the mesh of the finite element method and computational resources.

We look at the proposed algorithm using the example of the plane unidirectional source field in the scattering medium slab. In this case, the Peierls integral RTE (5.151) takes the form

$$\tilde{L}(\tau,\hat{\boldsymbol{l}}) = \begin{cases} \Delta(\tau,\hat{\boldsymbol{l}}) + \dfrac{\omega_0}{4\pi\mu}\displaystyle\int_0^{\tau} \mathrm{e}^{-(\tau-t)/\mu} \oint p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}')\tilde{L}(t,\hat{\boldsymbol{l}}')d\hat{\boldsymbol{l}}'dt, \quad \mu \geq 0; \\ \Delta(\tau,\hat{\boldsymbol{l}}) + \dfrac{\omega_0}{4\pi\,|\mu|}\displaystyle\int_{\tau}^{\tau_0} \mathrm{e}^{-(\tau-t)/\mu} \oint p(\hat{\boldsymbol{l}},\hat{\boldsymbol{l}}')\tilde{L}(t,\hat{\boldsymbol{l}}')\,d\hat{\boldsymbol{l}}'\,dt, \quad \mu < 0\,; \end{cases} \tag{5.154}$$

where the source function is:

$$\Delta(\tau,\hat{\boldsymbol{l}}) = \sum_{k=0}^{\infty} \frac{2k+1}{4\pi}\Delta_k(\tau,\mu)P_k(\hat{\boldsymbol{l}}\cdot\hat{\boldsymbol{l}}_0)\,,$$

$$\Delta_k(\tau,\mu) = \begin{cases} \dfrac{\mu_0\omega_0 p_k}{\mu_0 - d_k\mu}\left\{\exp\left[(\mu_0 - d_k\mu)\,\dfrac{\tau}{\mu\mu_0}\right] - 1\right\}\mathrm{e}^{-\tau/\mu} - \exp\left(-d_k\dfrac{\tau}{\mu_0}\right) + \exp(-\tau/\mu_0)\,, \\ \mu \geq 0\,; \\ \left(\dfrac{\mu_0\omega_0 p_k}{\mu_0 - d_k\mu}\left\{1 - \exp\left[(\mu_0 - d_k\mu)\,\dfrac{\tau_0 - \tau}{\mu\mu_0}\right]\right\} - 1\right)\exp\left(-d_k\dfrac{\tau}{\mu_0}\right) + \exp(-\tau/\mu_0)\,, \\ \mu < 0\,. \end{cases} \tag{5.155}$$

We used the set of the coaxial cylinders in the capacity of the mesh. The results of the radiance calculation of the reflected radiation are shown in Fig. 5.11 with the few iterations that are compared with the exact solution (a solid line) (5.69).

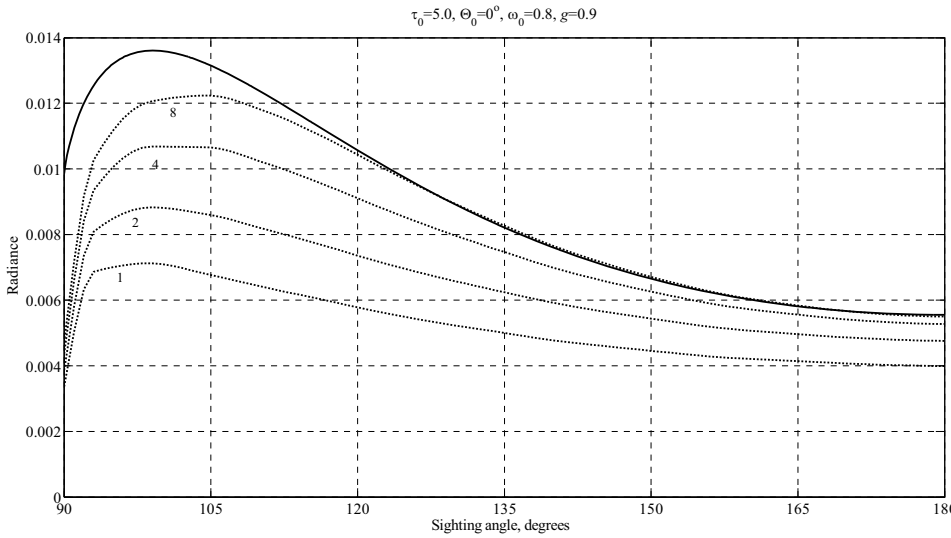


Fig. 5.11. Convergence of the method of iteration for the reflected radiation at $\tau_0 = 5.0$, $\theta_0 = 0$, $\omega_0 = 0.8$, $g = 0.9$. The solid curve is the exact solution. Numbers near the dotted curves give the number of the iteration.

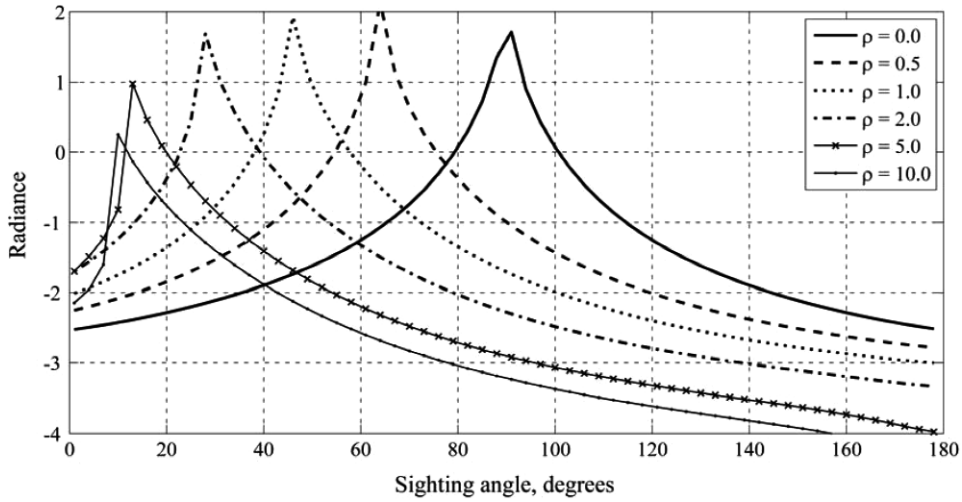


Fig. 5.12. The radiance angle distribution of the light field of the point isotropic source at the slab boundary. ρ is the distance to the point from the origin in the XOY-plane.

Note that the convergence for the downward radiation is essentially better and the amount of iterations does not exceed 2–4.

Now we go to the simplest 3-D boundary-value problem of a point isotropic source in the scattering medium slab. Using expressions (5.142) and (5.143), and the coaxial cylindrical mesh, we calculated RAD on one slab bound from a point isotropic source allocated on the other bound in the origin. The position of all the points in the space is specified by (τ, ρ) where ρ is the distance to the point from the origin in XOY-plane. The results are given in the Fig. 5.12 for the slab with optical parameters: $\omega_0 = 0.8$, $\tau = 1.0$ and $g = 0.9$.

Notice that in the calculation program only the source function $L_0(\boldsymbol{r}, \hat{\boldsymbol{l}})$ is changed while the rest of the program remains unchangeable. Generally speaking, the suggested approach of the anisotropic part of the solution elimination could be combined with any numerical method of the boundary method solution for the regular part.

5.11 Conclusion

From the analysis of the angular spectrum of the radiance distribution, it is possible to formulate the approximate equation describing an anisotropic part of the solution as the small-angle modification of the spherical harmonic method (MSH). This approach makes it possible to eliminate analytically the singularities from the RTE solution and to state the boundary-value problem for the regular part of the solution. The analytic form of MSH as a decomposition on the surface harmonic essentially simplifies the calculation of the source function in the equation for the regular part.

Within the bounds of MSH we could get the solution of RTE for all the fundamental sources that allows us to use this approach for eliminating the anisotropic

part of the solution in the arbitrary medium geometry. The role of the singularities of the radiance angular distribution essentially grows by the solution of the 3-D problems of the radiative transfer theory. In the conditions of the confined source presence or, what is equivalent, the sharp changes of the medium parameters in the case of the arbitrary three-dimensional geometry there are singularities in the radiance angular distribution, not only in the direct radiation, but also in the first two orders of scattering.

However, in the vectorial case there is still another problem: the reference planes of the incident and scattered rays as well as the scattering plane do not coincide. Therefore it is necessary to apply the rotator $\vec{R}(\chi)$ that disturbs the transformation symmetry of different Stokes parameters and makes it impossible to use the addition theorem for the surface harmonics. Kuščer and Ribarič (1959) proposed the polarization determination on the basis of the circular polarization, which is connected with Stokes polarization by the linear relation. In this case the rotator takes a simpler form that makes it possible to use for every polarization component the generalized spherical harmonic, for which its own form of the addition theorem is correct. Unfortunately in CP-presentation all the coefficients in VRTE become complex, which makes difficulties for using the effective numerical solution methods, which are based on the sorting algorithm. Therefore with the VRTE solution at first we convert equation to CP-presentation, subtract MSH, get the equation for the regular part of the solution, apply the addition theorem for generalized spherical harmonics, then we return to SP-presentation and finally solve the obtained equation by DOM.

The regular part of the solution is found numerically by the finite element method that enables us to extend the solution to the arbitrary medium geometry. The application of the developed method for the solution of the simple problem of the medium slab showed its high performance. However, its expansion to the case of the arbitrary medium geometry demands the development of the mesh-building technique for the numerical determination of the regular part of solution.

Acknowledgments

We would like to thank the participants in the seminar 'The photometric theory of the diffuse light field' of the Light Engineering Department of Moscow Power-Engineering Institute (TU) for the discussion of the results. We wish to express our gratitude to O. A. Popov for his invaluable help in translating this work into English. We are indebted to A. Kokhanovsky for lots of keen discussions, continuous encouragement and many useful comments on earlier versions of the manuscript. We express sincere gratitude to I. Laszlo for his kind assistance in our calculations with the code DISORT 2.0.

References

Apresyan, L.A., and Yu.A. Kravtsov, 1996: *Radiation Transfer: Statistical and Wave Aspects*, Amsterdam: Gordon & Breach.

Astakhov I.E., V.P. Budak, D.V. Lisitsin, V.A. Selivanov, 1994: Solution of the vector radiative transfer equation in the small angle approximation of the spherical harmonics method, *Atmospheric and Oceanic Optics*, **7**, 398–403.

Born, M., and E. Wolf, 1975: *Principles of Optics*, Oxford: Pergamon Press.

Bothe, W., 1929: Die Streuabsorption der Elektronenstrahlen, *Zeit.f. Physik*, **54**, 161–178.

Budak, V.P., and S.E. Sarmin, 1990: Solution of the radiative transfer equation by the method of spherical harmonics in the small-angle modification, *Atmospheric and Oceanic Optics*, **3**, 898–903.

Budak, V.P., and S.V. Korkin, 2008a: On the solution of a vectorial radiative transfer equation in an arbitrary three-dimensional turbid medium with anisotropic scattering, *J. Quant. Spectrosc. Radiat. Transf.*, **109**, 220–234.

Budak, V.P., and S.V. Korkin, 2008b: The aerosol influence upon the polarization state of the atmosphere solar radiation, *International Journal of Remote Sensing*, **29**, 2469–2506.

Budak, V.P., and S.V. Korkin, 2008c: The spatial polarization distribution over the dome of the sky for abnormal irradiance of the atmosphere, *J. Quant. Spectrosc. Radiat. Transf.*, **109**, 1347–1362.

Budak, V.P., K.V. Klimenko, D.A. Klyuykov, 2008: Calculation of light field created by point unidirectional source in 3-dimensional scattering media. *Current Problems in Atmospheric Radiation (IRS 2008): Proc. Int. Rad. Sym. (IRC/IAMAS), Foz do Igaucu, Brazil, 3–8 August 2008.*

Case, K.M., and P.F. Zweifel, 1967: *Linear Transport Theory*, Reading, MA: Addison-Wesley.

Chandrasekhar, S., 1950: *Radiative Transfer*, London: Oxford University Press.

Chwolson, O.D., 1889: Grundzüge einer matimatischen Theorie der inneren Diffusion des Licht, *Melanges Phys. et Chim. Tires du Bull. De l'Acad. Imp. de Sci. de St.-Petersb.*, **13**, 83.

Deirmendjian, D., 1969: *Electromagnetic Scattering of Spherical Polydispersions*, New York: American Elsevier.

Dolin, L.S., 1964: Beam Description of weakly-inhomogeneous wave fields, *Izv. VUZOV. Radiofizika*, **7**, 559 (in Russian).

Eddington, A.S., 1916: On the radiative equilibrium of the stars, *Mon. Not. R. Astr. Soc.*, **76**, 16–35.

Gelfand, I.M., R.A. Minlos, Z.Ya. Shapiro, 1963: *Representations of the Rotation and Lorentz Groups and their Applications*, Oxford: Pergamon Press.

Germogenova, T.A., 1986: *Local Properties of the Transport Equation Solutions*, Moscow: Nauka (in Russian).

Goudsmit, S., and J.L. Saunderson, 1940: Multiple scattering of electrons, *Phys. Rev.*, **57**, 24–29.

Gratton, L., 1939: Problemi di equilibrio radiativo e loro soluzione mediante I polinomi di Legendre, *Memorie della Societa Astronomia Italiana*, **10**, 309–325.

Hovenier, J.W., C. van der Mee, H. Domke, 2004: *Transfer of Polarized Light in Planetary Atmosphere. Basic Concepts and Practical Methods*, Dordrecht: Kluwer Academic Publishers.

Irvine, W.M., 1968: Diffuse reflection and transmission by cloud and dust layers, *J. Quant. Spectrosc. Radiat. Transf.*, **8**, 471–485.

Jeans, J.H., 1917: The equations of radiative transfer of energy, *Mon. Not. R. Astr. Soc.*, **78**, 28–36.

Karp, A.H., J. Greenstadt, J.A. Fillmore, 1980: Radiative transfer through an arbitrary thick scattering atmosphere, *J. Quant. Spectrosc. Radiat. Transf.*, **24**, 391–406.

Koch, R., and R. Becker, 2004: Evaluation of quadrature schemes for the discrete ordinates method, *J. Quant. Spectrosc. Radiat. Transf.*, **84**, 423–435.

Kompaneets, A.S., 1947: Multiple scattering of pencil beam of fast electrons, *Zhurnal eksperimentalnoy i teoreticheskoy fiziki* (*Journal of Experimental and Theoretical Physics*), **17**, 1059–1062 (in Russian).

Krylov, V.I., 2006: *Approximate Calculation of Integrals*, New York: Dover Publications.

Kuščer, I., and M. Ribarič: 1959, Matrix formalism in the theory of diffusion of light, *Optica Acta*, **6**, 42–51.

Lenoble, J., 1985: *Radiative Transfer in Scattering and Absorbing Atmospheres: Standard Computational Procedures*, Hampton, VA: A. Deepak.

Lommel, E., 1889: Die Photometric der diffusen Zuruckwerfung, *Sitzber. Acad. Wissensch. Munchen*, **17**, 95–124.

Lorentz, A.H., 1916: *The Theory of Electrons and its Applications to the Phenomena of Light and Radiant Heat*, Leipzig: B.G. Teubner.

Marchuk, G.I., G.A. Mikhailov, M.A. Nazaraliev, R.A. Darbinjan, B.A. Kargin, B.S. Elepov, 1980: *The Monte Carlo Methods in Atmospheric Optics*. Springer Series in Optical Sciences, Berlin: Springer-Verlag.

Mark, C., 1947: The neutron density near a plane surface, *Phys. Rev.*, **72**, 558–564.

Milne, E.A., 1926: The reflection effect in eclipse binaries, *Mon. Not. R. Astr. Soc.*, **87**, 43–55.

Mishchenko, M.I., 2008: Multiple scattering, radiative transfer, and weak localization in discrete random media: unified microphysical approach, *Rev. Geophys.*, **46**, RG2003.

Mitchell, A.R., and R. Wait, 1976: *The Finite Element Method in Partial Differential Equation*. London: Wiley.

Mobley C.D., L.K. Sundman, E. Boss, 2002; Phase function effects on oceanic light fields, *Appl. Opt.*, **41**, 1035–1050.

Molière, G., 1948: Theorie der Streuung schneller geladener Teilchen II. Mehrfach- und Vielfachstreuung, *Zeit. f. Natur.*, **3a**, 78–97.

Muehlschlegel, B., and H. Koppe, 1958: Theorie der Vielfachstreuung polarisierter Elektronen, *Zeit. f. Physik*, **150**, 474–496.

Nakajima, T., M. Tanaka, 1986: Matrix formulations for the transfer of solar radiation in a plane-parallel scattering atmosphere, *J. Quant. Spectrosc. Radiat. Transf.*, **35**, 13–21.

Plass, G.N., G.W. Kattawar, F.E. Catchings, 1973: Matrix operator theory of radiative transfer. 1: Rayleigh scattering, *Appl. Opt.*, **12**, 314–329.

Romanova, L.M., 1962a: Small angle approximation of radiative transfer equation solution ant its improvement, *Izv. Akademii Nauk Seriya Fizicheskaya* (Proc. Academy of Science USSR. Geophys. Series), **8**, 1108–1112 (in Russian).

Romanova, L.M., 1962b: Solution of the radiative transfer equation in the case of scattering phase function greatly differs from a spherical one., *Opt. Spectrosc.*, **13**, 429–435 (in Russian).

Rozanov, A., V. Rozanov, M. Buchwitz, A. Kokhanovsky, J.P. Burrows, 2005: SCIATRAN 2.0 – A new radiative transfer model for geophysical applications in the 175–2400 nm spectral region, *Advances in Space Research*, **36**, 1015–1019.

Rozenberg, G.V., 1977: Light ray (On the theory of light field), *Uspekhi Fiz. Nauk* (Advances in Physical Sciences), **121**, 97–138 (in Russian).

Siewert, C.E., 2000: A discrete-ordinates solution for radiative-transfer models that include polarization effects, *J. Quant. Spectrosc. Radiat. Transf.*, **64**, 227–254.

Snyder, H.S., and W.T. Scott, 1949: Multiple scattering of fast charged particles, *Phys. Rev.*, **76**, 220–225.

Stamnes, K., and R.A. Swanson, 1981: A new look at the discrete ordinate method for radiative transfer calculation in anisotropicaly scattering atmosphere, *J. Atm. Sci.*, **38**, 387–399.

Stokes, G.G., 1862: On the intensity of the light reflected from or transmitted through a pile of plates, *Proc. R. Soc. London*, **11**, 545–556.

Sushkevich, T.A., 2005: *Mathematical Model of Radiative Transfer*, Moscow: BINOM. Laboratoria znaniy (in Russian).

Sykes, J.B., 1951: Approximate integration of the equation of transfer, *Month. Not. R. Astroph. Soc.*, **111**, 378–386.

Thomas, G.E., and K. Stamnes, 2002: *Radiative Transfer in the Atmosphere and Ocean*, Cambridge: Cambridge University Press.

Umov, N., 1905: Chromatische Depolarisation durch Lichtzerstreuung, *Phys. Zeit.*, **6**, 674–676.

Ustinov, E.A., 1988: Spherical harmonics method: application to the calculation of the polarized light transfer in the vertical inhomogeneous planet atmosphere, *Mathematical Tool, Kosmicheskie Issledovanya (Space Research)*, **26**, 550–562 (in Russian).

Van de Hulst, H.C., 1948: Scattering in planetary atmosphere, *Ap. J.*, **107**, 220–246.

Van de Hulst, H.C., 1957: *Light Scattering by Small Particles*, New York: J.W.&S.

Vilenkin, N. Ya., 1965: *Special Functions and Theory of Group Representations*, Moscow: Izdat. 'Nauka' (Transl. Math. Monographs, 1968, 22, Amer. Math. Soc, Providence, RI).

Wang, M.C., and E. Guth, 1951: On the theory of multiple scattering, particulary of charged particles, *Phys. Rev.*, **84**, 1092–1111.

Wentzel, G., 1922: Zur theorie der Streuung von β-Strahlen, *Ann. d. Phys.*, **69**, 335–368.

Wick, G.C., 1943: Über ebene Diffussionsprobleme, *Zeit. f. Phys.*, **121**, 702–718.

Wiscombe, W.J., 1977: The Delta–M method: rapid yet accurate radiative flux calculations for strongly asymmetric phase functions. *J. Atm. Sci.*, **34**, 1408–1422.

Zege, E.P., and L.I. Chaikovskaya, 1981: Expansion of scattering matrix elements of sea water in terms of generalized spherical function, *Izv. AN SSSR. Fizika atmosfery i okeana (Proceedings of the Academy of Sciences of USSR. Physics of Atmosphere and Ocean)*, **17**, 604–613 (in Russian).

Zege, E.P., A.P. Ivanov, I.L. Katsev, 1991: *Image transfer through a scattering medium*, Heidelberg: Springer Verlag.

6 Code SHARM: fast and accurate radiative transfer over spatially variable anisotropic surfaces

Alexei Lyapustin, Tolegen Muldashev and Yujie Wang

Numerous numerical methods have been developed to solve a plane-parallel radiative transfer problem, including discrete ordinates (Stamnes et al., 1988; Spurr et al., 2001), spherical harmonics (Dave, 1975; Benassi et al., 1984), adding-doubling (Hansen and Hovenier, 1971; Twomey, 1985), successive orders of scattering (Lenoble et al., 2007) etc. An extended reference to numerical methods and publicly available codes can be found in Lenoble (1985), Ricchiazzi et al. (1998), Mayer and Kylling (2005) and Cahalan (2005). In this chapter, we describe the method of spherical harmonics (MSH), in particular its very efficient implementation developed by Karp et al. (1980) and, later, Muldashev et al. (1999). From a numerical standpoint, here are several main components of the spherical harmonics solution: (1) obtaining the system of linear differential equations of MSH, (2) its reduction to the system of linear algebraic equations using singular value decomposition (SVD), (3) use of a system's matrix symmetry to halve its size for SVD transformation with $\sim$8 times gain in speed, and finally, (4) angular smoothing of the solution for radiance calculations in arbitrary directions. The detail of MSH for the 1-D radiative transfer problem with a uniform surface, and for the 3-D problem with a spatially variable surface, are presented in sections 6.1 and 6.5, respectively. Section 6.2 provides an overview of the 1-D radiative transfer code SHARM (Muldashev et al., 1999; Lyapustin and Wang, 2005) which is one of the most numerically efficient scalar codes. SHARM performs simultaneous monochromatic calculations for multiple sun-view geometries, and allows the user to make multi-wavelength calculations in one run. The code is user-friendly, featuring built-in aerosol models and the most popular models of the bi-directional reflectance factor (BRF) of land and wind-ruffled water surface. Comparisons of SHARM with the benchmark code DISORT showed agreement to 0.02%.

With all its efficiency, a numerical solution cannot be used for atmospheric correction of large volumes of space-borne data. The spectral surface bi-directional reflectance and albedo are primary inputs for many higher level products of the land processing, which is an important component of the Earth Observing System (EOS) (King and Greenstone, 1999). Accurate and operationally viable atmospheric correction requires an analytically explicit approach to derive the parameters of a surface BRF model. In section 6.3, we use the Green's function method to obtain an accurate semi-analytical solution for the top-of-atmosphere (TOA) radiance,

which explicitly depends on parameters of the linear kernel BRF model (Roujean et al., 1992). This solution is used to build very efficient algorithms of aerosol retrievals and atmospheric correction from the EOS Moderate Resolution Imaging Spectroradiometer (MODIS).

Land surfaces are not homogeneous. Variability of surface brightness gives rise to the horizontal fluxes of radiation in the atmosphere directed from the brighter to the darker areas, which is commonly called the 'adjacency' effect. This 3-D effect becomes more important with the increase in the resolution of space-borne imaging systems. An analytical 3-D Green's function solution for the TOA radiance over a variable surface with anisotropic reflectance is derived in section 6.4. Practical application of the derived formula relies on the knowledge of the 3-D atmospheric Green's function, or its Fourier transform called the optical transfer function (OTF). The MSH method of computing the OTF is described in section 6.5. These developments led to the creation of code SHARM-3D described in section 6, which is designed for radiance computations over spatially variable anisotropic surfaces with laterally homogeneous atmosphere. SHARM-3D uses accurate parameterizations of the Green's function solution and a pre-computed look-up table (LUT) of OTF which makes it very fast. Comparisons between SHARM-3D and the rigorous 3-D code SHDOM (Evans, 1998) showed an agreement to within 0.5–1%. At the same time, SHARM-3D is at least a thousand times faster and allows the processing of large scenes with the side dimension of 10^3–10^4 on an ordinary laptop, whereas SHDOM's memory requirements limit the size to about 32 to 64 pixels.

6.1 The method of spherical harmonics: homogeneous surface

Let us consider a plane-parallel radiative transfer problem with a uniform surface. Scattering and absorbing properties of the atmosphere, described by the optical thickness $0 \le \tau \le \tau_0$, single scattering albedo $\omega(\tau)$, and scattering function $\chi(\gamma, \tau)$, where γ is the scattering angle, may vary vertically. Directions are defined in the spherical coordinate system $(\tau,\ \mu = \cos\theta,\ \varphi)$ where θ and φ are zenith and azimuthal angles. The polar axis τ starts at the top of atmosphere (TOA), so $\mu > 0$ for the downward directions and $\mu < 0$ for the upward directions. The intensity of diffuse radiation $I(\tau;\ \mu,\ \varphi)$ is a solution of the following boundary-value problem:

$$\mu \frac{\partial I(\tau,\mu,\varphi)}{\partial \tau} + I(\tau,\mu,\varphi) = \frac{\omega}{4\pi}\int_0^{2\pi} d\varphi' \int_{-1}^{1} \chi(\tau,\gamma) I(\tau,\mu',\varphi') d\mu' + S_\lambda \frac{\omega}{4}\chi(\tau,\gamma_0)\,\exp\left(-\frac{\tau}{\mu_0}\right) \tag{6.1}$$

$$I(0;\mu,\varphi) = 0,\ \mu > 0 \tag{6.1a}$$

$$I(\tau_0,\mu,\varphi) = S_\lambda \mu_0 \rho(\mu_0,\mu,\varphi-\varphi_0) e^{-\tau_0/\mu_0} + \frac{1}{\pi}\int_0^{2\pi} d\varphi' \int_0^1 \rho(\mu',\mu,\varphi-\varphi') I(\tau_0,\mu',\varphi')\mu' d\mu',\ \mu < 0 \tag{6.1b}$$

Here, ρ is a unitless surface BRF and πS_λ is the extraterrestrial spectral solar irradiance. The scattering function is normalized to unity, $\frac{1}{2}\int_0^\pi \chi(\gamma)\,\sin\gamma\,d\gamma = 1$. To solve problem (6.1), the scattering function is expanded into a Legendre polynomial series,

$$x(\gamma) = \sum_{n=0}^{L} \chi_n P_n(\cos\gamma), \quad \chi_k = \frac{2k+1}{2}\int_{-1}^{1} \chi(y)P_k(y)\,dy. \tag{6.2}$$

Application of the addition theorem gives an azimuthal series for the phase function,

$$\begin{aligned} \chi(\tau,\gamma) &= (2-\delta_{0,m})\sum_{m=0}^{\infty}\chi^m(\tau,\mu,\mu')\,\cos m(\varphi-\varphi'), \\ \chi^m(\tau,\mu,\mu') &= \sum_{k=m}^{\infty}\chi_k(\tau)Y_k^m(\mu)Y_k^m(\mu'), \end{aligned} \tag{6.3}$$

where $Y_k^m(\mu)$ are associated Legendre polynomials, and $\delta_{0,m}$ is the Kronecker symbol. By analogy, radiance and surface bi-directional reflectance are expanded into a Fourier series:

$$\begin{aligned} I(\tau,\mu,\varphi) &= (2-\delta_{0,m})\sum_{m=0}^{\infty} I^m(\tau,\mu)\,\cos m\varphi, \\ \rho(\mu_0,\mu,\varphi) &= (2-\delta_{0,m})\sum_{m=0}^{\infty}\rho^m(\mu_0,\mu)\,\cos m\varphi. \end{aligned} \tag{6.4}$$

Substitution of (6.3), (6.4) into (6.1) gives the problem for the mth Fourier component of intensity:

$$\begin{aligned} \mu\frac{\partial I^m(\tau,\mu)}{\partial\tau} + I^m(\tau,\mu) &= \frac{1}{2}\int_{-1}^{1}\chi^m(\tau,\mu',\mu)I^m(\tau,\mu')\,d\mu' \\ &\quad + \frac{1}{4}S_\lambda\chi^m(\tau,\mu_0,\mu)\,e^{-\tau/\mu_0}, \end{aligned} \tag{6.5}$$

$$I^m(0,\mu) = 0,\ \mu > 0, \tag{6.5a}$$

$$\begin{aligned} I^m(\tau_0,\mu) &= S_\lambda\mu_0\rho^m(\mu_0,\mu)e^{\tau_0/\mu_0} \\ &\quad + 2\int_0^1 \rho^m(\mu',\mu)\mu' I^m(\tau_0,\mu')\,d\mu',\ \mu < 0. \end{aligned} \tag{6.5b}$$

Next, $I^m(\tau,\mu)$ is expanded into Legendre series:

$$I^m(\tau,\mu) = \sum_{k=m}^{L_m}\frac{2k+1}{2}\varphi_k^m(\tau)Y_k^m(\mu), \quad \varphi_k^m(\tau) = \int_{-1}^{1} I^m(\tau,\mu)Y_k^m(\mu)\,d\mu, \tag{6.6}$$

where the choice of parameter L_m will be explained later. Using expansions (6.6), (6.3) and equation (6.5), one can obtain a system of linear differential equations for the moments of intensity $\varphi_k^m(\tau)$. To this end, equation (5) is multiplied by $Y_k^m(\mu)$

for $k = m,\ m+1, \ldots, L_m$ and integrated over μ from -1 to 1. This leads to the system of differential equations, which can be written in matrix form:

$$A^m \frac{d\vec{\varphi}^m(\tau)}{d\tau} + C^m \vec{\varphi}^m(\tau) = e^{\tau/\mu_0} \vec{f}^m(\tau), \tag{6.7}$$

where A^m is a symmetrical matrix of the order $(L_m - m + 1)$ with non-zero off-diagonal elements,

$$A^m = \begin{pmatrix} 0 & a_1 & 0 & & 0 \\ a_1 & 0 & a_2 & & \vdots \\ 0 & a_2 & 0 & & \\ & & & \ddots & a_{L_m - m} \\ 0 & \cdots & & a_{L_m - m} & 0 \end{pmatrix}, \quad a_k = \sqrt{k(2m+k)},$$

and C^m is a diagonal matrix, $C^m = \mathrm{diag}[2m + 1 - \chi_m(\tau),\ 2m + 3 - \chi_{m+1}(\tau), \ldots, 2L_m + 1 - \chi_{L_m}(\tau)]$.

Vectors $\vec{\varphi}^m(\tau)$, $\vec{f}^m(\tau)$ have the form $\vec{f}^m(\tau) = [f^m_m(\tau), f^m_{m+1}(\tau), \ldots, f^m_{L_m}(\tau)]^T$, and elements of vector $\vec{f}^m(\tau)$ are equal to $f^m_{m+k}(\tau) = \frac{S_\lambda}{2} \chi_{m+k}(\tau) Y^m_{m+k}(\mu_0)$. To conclude the system, we set $\varphi^m_{L_m+1}(\tau) = 0$ in the last equation of system (6.7). Parameter L_m is selected such that the order of system (6.7) is even to ensure continuity of vector $\vec{\varphi}^m(\tau)$ on the boundaries of atmospheric layers. In particular, if L_0 is odd, then $L_m = L_0 + \delta_m$, where $\delta_m = 0$ if m is even, and $\delta_m = 1$ if m is odd.

We use Marshak's form of the boundary conditions (Marshak, 1947; Dave, 1975). For example, the lower boundary condition is written as follows (Lyapustin and Muldashev, 2000):

$$(I - 2N^m_{od})\vec{\varphi}^m_{od}(\tau_0) + (G^m - 2N^m_{ev})\vec{\varphi}^m_{ev}(\tau_0) = 2S_\lambda e^{\tau_0/\mu_0} \vec{Q}^m(\mu_0). \tag{6.8}$$

Here, I is an identity matrix, $\vec{\varphi}^m_{od}(\tau_0)$ and $\vec{\varphi}^m_{ev}(\tau_0)$ are vectors consisting of the odd and even elements of vector $\vec{\varphi}^m(\tau_0)$, and vector $\vec{Q}^m(\mu')$ is defined as follows:

$$Q^m_l(\mu') = \int_{-1}^{0} \mu' \rho^m(\mu', \mu) Y^m_{m-\delta_m+2l-1}(\mu)\, d\mu. \tag{6.9a}$$

G^m, N^m_{od}, N^m_{ev} are quadratic matrices with elements

$$G^m_{l,k} = -(4k-3) \int_0^1 Y^m_{m-\delta_m+2l-1}(\mu) Y^m_{m+\delta_m+2k-2}(\mu)\, d\mu, \tag{6.9b}$$

$$(N^m_{ev})_{l,p} = (4p + 2m + 2\delta_m - 3) \int_0^1 Q^m_l(\mu') Y^m_{m+\delta_m+2p-2}(\mu')\, d\mu', \tag{6.9c}$$

$$(N^m_{od})_{l,p} = (4p + 2m - 2\delta_m - 1) \int_0^1 Q^m_l(\mu') Y^m_{m-\delta_m+2p-1}(\mu')\, d\mu',$$

$$l,\ p = 1, 2, \ldots, \frac{L_m - m + 1}{2}. \tag{6.9d}$$

Matrix G is calculated using recurrence formulas of Dave (1975), while vector Q and matrices N^m_{od}, N^m_{ev} are computed numerically using Gaussian quadrature.

With described boundary conditions, the problem for vector $\vec{\varphi}^m(\tau)$ becomes:

$$A^m \frac{d\vec{\varphi}^m(\tau)}{d\tau} + C^m \vec{\varphi}^m(\tau) = e^{\tau/\mu_0} \vec{f}^m(\tau), \tag{6.10}$$

$$\vec{\varphi}^m_{od}(0) - G^m \vec{\varphi}^m_{ev}(0) = 0, \tag{6.10a}$$

$$(I - 2N^m_{od})\vec{\varphi}^m_{od}(\tau_0) + (G^m - 2N^m_{ev})\vec{\varphi}^m_{ev}(\tau_0) = 2S_\lambda e^{-\tau_0/\mu_0}\vec{Q}^m(\mu_0). \tag{6.10b}$$

When surface is Lambertian, problem (6.1) splits into two problems: one is for the path radiance $D(\tau, \mu, \varphi)$ with zero boundary conditions and the other one for a diffuse atmospheric transmittance (Muldashev et al., 1998). In many applied problems, including atmospheric correction, it is convenient to have these functions separately along with other radiative transfer characteristics including fluxes and spherical albedo of the atmosphere. Below, we describe solution for path radiance, which is a particular case of problem (6.10) with zero matrices N^m_{od}, N^m_{ev} and zero vector $\vec{Q}^m(\mu_0)$.

6.1.1 Solution for path radiance

One of the most efficient solutions of problem (6.10) was developed by Karp et al.(1980). Using singular value decomposition (SVD), the system of linear differential equations of MSH is initially reduced to the system of linear algebraic equations. Further, due to its symmetry, the size of the system's matrix is effectively halved, reducing computing time by a factor of ∼8.

Let the atmosphere consist of homogeneous layers, each of an arbitrary optical thickness $\Delta\tau_i = \tau_i - \tau_{i-1}$, $i = 1, \ldots, n$, and a constant matrix $C(\tau) = C_i$. Since the problems for different m are solved independently, index m is omitted in following discussion. Integration of Eq. (6.10) within the layer $[\tau_{i-1}, \tau_i]$ gives an analytical solution at the interfaces:

$$e^{B_i \Delta\tau_i}\vec{\varphi}(\tau_i) - \vec{\varphi}(\tau_{i-1}) = \beta_i, \quad B_i = A^{-1}C_i, \ \beta_i = \int_{\tau_{i-1}}^{\tau_i} e^{B_i(t-\tau_{i-1})}A^{-1}e^{-\frac{t}{\mu_0}}\vec{f}(t)\,dt. \tag{6.11}$$

To comply with boundary conditions, a unitary transformation P_0 is defined which sorts a vector into the odd and even parts,

$$P_0\vec{\varphi} = \begin{pmatrix} \varphi_{ev} \\ \varphi_{od} \end{pmatrix} = \vec{g} = \begin{pmatrix} g^t \\ g^b \end{pmatrix}.$$

Multiplication of (6.11) by P_0 gives the following problem:

$$-\vec{g}_{i-1} + P_0 e^{B_i \Delta\tau_i} P_0^T \vec{g}_i = P_0\beta_i, \ i = 1, 2, \ldots, n \tag{6.12}$$

$$\vec{g}^{\,b}_0 = G\vec{g}^{\,t}_0, \quad \vec{g}^{\,b}_n = -G\vec{g}^{\,t}_n. \tag{6.12a}$$

The next step is to find matrix exponential $e^{B_i\Delta\tau_i}$. One efficient way is to use a singular value decomposition of matrix B. When $\omega < 1$, B has different eigenvalues

and a complete set of linearly independent eigenvectors. In this case, it can be written as a product

$$B = U \begin{pmatrix} \Lambda_+ & 0 \\ 0 & \Lambda_- \end{pmatrix} U^{-1},$$

where $\Lambda_\pm$ are diagonal matrices, $\Lambda_+ = \mathrm{diag}\{\lambda_1, \ldots, \lambda_{(L_m - m+1)/2}\}$, $\Lambda_- = -\Lambda_+$, and $\pm\lambda_k$ are the eigenvalues of matrix B, and U is a matrix of eigenvectors. When $\omega = 1$, B has two zero eigenvalues $+\lambda_1 = -\lambda_1 = 0$, which are associated with only one eigenvector. In this case, matrix B can be reduced to the Jordan form

$$B = UJU^{-1}, \quad J = \begin{pmatrix} \Lambda_+ & R \\ 0 & \Lambda_- \end{pmatrix}, \tag{6.13}$$

where the missing eigenvector is replaced by a principal vector (defined as a vector which satisfies $(B - \lambda I)^2 \vec{x} = 0$) (Karp et al., 1980). All elements of matrix R are equal to zero, except for $R_{11} = r$. When $\omega \to 1$ then $\lambda_1 \to 0$, and eigenvectors $\vec{u}_1$ and $\vec{u}_2$ corresponding to $+\lambda_1$ and $-\lambda_1$ become nearly parallel. This results in poor conditioning of matrix U and in unstable matrix inversion. Replacing $\vec{u}_2 = \vec{u}_1 - \vec{u}_2$ corrects matrix U, and preserves decomposition (6.13) for matrix B. Thus, the Jordan form provides a uniform decomposition of matrix B and its continuous dependence on parameter ω.

Now, it is easy to exponentiate matrix B:

$$e^{B\Delta\tau} = Ue^{J\Delta\tau}U^{-1}, \tag{6.14}$$

where

$$e^{J\Delta\tau} = \begin{pmatrix} e^{\Lambda_+\Delta\tau} & T \\ 0 & e^{\Lambda_-\Delta\tau} \end{pmatrix},$$

and

$$T_{11} = \frac{r}{2\lambda_1}(e^{\lambda_1\Delta\tau} - e^{-\lambda_1\Delta\tau}),\ T_{11} = r\Delta\tau \text{ for } \lambda_1 = 0.$$

The other elements of matrix T are equal to zero. Substituting (6.14) into (6.12) and multiplying the result by $SU^{-1}P_0^T$, where S is a scaling matrix,

$$S = \begin{pmatrix} 0 & I \\ e^{\Lambda_-\Delta\tau} & 0 \end{pmatrix},$$

we finally obtain

$$-S_iW_i\vec{g}_{i-1} + H_iW_i\vec{g}_i = \vec{d}_i,\ i = 1, 2, \ldots, n, \tag{6.15}$$

where $W = (P_0U)^{-1}$, and

$$H = Se^{J\Delta\tau} = \begin{pmatrix} 0 & I \\ e^{\Lambda_-\Delta\tau} & 0 \end{pmatrix} \begin{pmatrix} e^{\Lambda_+\Delta\tau} & T \\ 0 & e^{\Lambda_-\Delta\tau} \end{pmatrix} = \begin{pmatrix} 0 & e^{\Lambda_-\Delta\tau} \\ I & e^{\Lambda_-\Delta\tau}T \end{pmatrix}.$$

The last equation shows that the scaling transformation (Karp et al., 1980) eliminates positive exponents in the coefficients of system (6.15). Along with the boundary conditions (6.12a), equation (6.15) represents a linear system of equations with a block matrix which is solved by a Gauss elimination method. The problem (6.10) with anisotropic surface reflectance differs only in the low boundary condition and is solved similarly.

6.1.2 Correction function of MSH

Once moments are found, the mth Fourier harmonic of intensity can be found from

$$u^m(\tau,\mu)=\sum_{k=m}^{L_m}\frac{2k+1}{2}\varphi_k^m(\tau)Y_k^m(\mu).$$

For arbitrary μ, this solution has an oscillating error around the exact value due to truncation of the scattering function, reduction of the order of MSH system, and due to an approximate form of boundary conditions. The amplitude of oscillations is largest at the boundaries of layers. It decreases with the increase of the order of approximation L, yet the computing time grows as $\sim L^3$. For this reason, special smoothing procedures are applied as a tradeoff between the speed of computations and the desired accuracy.

One of the most precise smoothing methods is integration of the source function (ISF) (Dave and Armstrong, 1974; Karp, 1981), where computed moments $\varphi_k^m(\tau)$ are only used to find the source function. Since this source function is integrated with exact boundary conditions and since the scattering function in the single scattering term can be represented accurately, the ISF solution is smooth and much more accurate than function $u^m(\tau,\mu)$. With its high accuracy, the ISF method is rather time-consuming. Even in a form of Karp (1981) that avoids numeric integration, it needs matrix inversion for each atmospheric layer and for each direction μ.

Muldashev et al. (1999) developed a 'correction function' method which preserves the accuracy of ISF but is considerably faster. Let us obtain a radiative transfer equation for an approximate solution $u^m(\tau,\mu)$. To this end, we multiply each kth equation of system (6.7) by $Y_{m+k}^m(\mu)/2$, $k=0,1,\ldots,L_m-m$, and sum up all equations with the following result:

$$\begin{aligned}\mu\frac{\partial u^m(\tau,\mu)}{\partial\tau}+u^m(\tau,\mu) \;&=\; \frac{1}{2}\int_{-1}^{1}\chi^m(\tau,\mu',\mu)u^m(\tau,\mu')d\mu'+\frac{1}{4}S_\lambda\chi^m(\tau,\mu_0,\mu)e^{\tau/\mu_0}\\ &+\frac{1}{2}\sqrt{(L_m-m+1)(L_m+m+1)}Y_{L_m+1}^m(\mu)\frac{d\varphi_{L_m}^m(\tau)}{d\tau}.\end{aligned}\tag{6.16}$$

If $I^m(\tau,\mu)$ is an exact solution, then $w^m(\tau,\mu)=u^m(\tau,\mu)-I^m(\tau,\mu)$ is the error of approximate solution (correction function). Let us rewrite (6.16) for function $w^m(\tau,\mu)$ with the boundary conditions:

$$\begin{aligned}\mu\frac{\partial w^m(\tau,\mu)}{\partial\tau}+w^m(\tau,\mu) \;&=\; \frac{1}{2}\int_{-1}^{1}\chi^m(\tau,\mu,\mu')w^m(\tau,\mu')\,d\mu'\\ &+F_1^m(\tau,\mu)+F_2^m(\tau,\mu),\end{aligned}\tag{6.17}$$

$$w^m(0,\mu)=u^m(0,\mu),\ \mu>0;\quad w^m(\tau_0,\mu)=u^m(\tau_0,\mu),\ \mu<0.\tag{6.17a}$$

Here, $F_1^m(\tau,\mu)=\frac{1}{4}S_\lambda\{\chi^m(\tau,\mu_0,\mu)-p^m(\tau,\mu_0,\mu)\}e^{-\tau/\mu_0}$ is an error of expansion of the right hand side of equation (6.5), where $p^m(\tau,\mu_0,\mu)$ stands for the exact mth azimuthal harmonic of the phase function;

$$F_2^m(\tau,\mu)=\frac{1}{2}\sqrt{(L_m-m+1)(L_m+m+1)}Y^m_{L_m+1}(\mu)\frac{d\varphi^m_{L_m}(\tau)}{d\tau}$$

is an expansion error of derivative $\mu[\partial u^m(\tau,\mu)]/\partial\tau$, and $u^m(0,\mu)$, $u^m(\tau_0,\mu)$ are the inaccuracies of Marshak's boundary conditions. The main error comes from truncations of the expansion series and appears in the single scattering term. Higher orders of scattering are considerably smoothed in angle. For this reason, it is sufficient to solve (6.17) in the single scattering approximation. For simplicity, we will write the solution for a homogeneous atmosphere:

$$\begin{aligned} w_1^m(\tau,\mu) &= e^{-\tau/\mu}u^m(0,\mu) \\ &+ \frac{1}{4}S_\lambda\frac{\mu_0}{\mu_0-\mu}(e^{-\tau/\mu_0}-e^{-\tau/\mu})[\chi^m(\tau,\mu_0,\mu)-p^m(\tau,\mu_0,\mu)] \\ &+ \sqrt{(L_m-m+1)(L_m+m+1)}\frac{Y^m_{L_m+1}(\mu)}{2\mu} \\ &\times\left\{\varphi^m_{L_m}(\tau)-e^{-\frac{\tau}{\mu}}\varphi^m_{L_m}(0)-\frac{1}{\mu}\int_0^\tau e^{(t-\tau)/\mu}\varphi^m_{L_m}(t)\,dt\right\},\ \mu>0, \end{aligned} \tag{6.18}$$

and

$$\begin{aligned} w_1^m(\tau,\mu) &= e^{\tau_0-\tau/\mu}u^m(\tau_0,\mu) \\ &+ \frac{1}{4}S_\lambda\frac{\mu_0}{\mu_0-\mu}(e^{-\tau/\mu_0}-e^{\tau_0-\tau/\mu}e^{-\tau_0/\mu_0})[\chi^m(\tau,\mu_0,\mu)-p^m(\tau,\mu_0,\mu)] \\ &+ \sqrt{(L_m-m+1)(L_m+m+1)}\frac{Y^m_{L_m+1}(\mu)}{2\mu} \\ &\times\left\{\varphi^m_{L_m}(\tau)-e^{\tau_0-\tau/\mu}\varphi^m_{L_m}(\tau_0)-\frac{1}{\mu}\int_{\tau_0}^\tau e^{(t-\tau)/\mu}\varphi^m_{L_m}(t)\,dt\right\},\ \mu<0. \end{aligned} \tag{6.19}$$

The integrals in (6.18)–(6.19) can be computed analytically using recurrent formula given by Muldashev et al. (1999). The final solution is a sum of an approximate solution and correction function:

$$\tilde{u}^m(\tau,\mu)=u^m(\tau,\mu)-w_1^m(\tau,\mu). \tag{6.20}$$

Such a correction using $w_1^m(\tau,\mu)$ gives an exact contribution of the single scattering regardless of the order of approximation of MSH. For a uniform atmosphere, this method is twice as fast as the ISF method. This advantage increases in computations for multiple-layer atmospheres and multiple sun-view geometries.

6.2 Code SHARM

The described MSH solution was implemented in code SHARM (Lyapustin and Wang, 2005) which solves the monochromatic unpolarized plane-parallel 1-D problem with vertically non-uniform atmosphere and several broadly used models of the land/ocean surface bi-directional reflectance. This code (as well as 3-D code

SHARM-3D) is available at http://neptune.gsfc.nasa.gov/bsb/subpages/index.php?section=Projects&content=SHARM. SHARM computes radiance and fluxes at the interfaces of atmospheric layers in the shortwave spectral region (0.4–2.5 μm). When the surface is Lambertian, it computes path radiance, upward atmospheric transmittance, spherical albedo of the atmosphere and the total radiance separately. SHARM performs simultaneous calculations for different solar zenith angles (SZA), view zenith angles (VZA), and view azimuths, and allows the user to handle multiple wavelengths in one run. Molecular scattering is added automatically according to the wavelength, surface height above the sea level, and selected atmospheric profile. The molecular absorption for each atmospheric layer should be provided by the user.

The code comes in a package with program *Phase*, which computes Legendre expansion coefficients of aerosol/cloud phase function, and automatically normalizes it. The rational spline interpolation is used to compute the phase function in the quadrature angles required for the Legendre expansion, and in the directions required for the single scattering calculations. The rational spline provides an accurate smooth interpolation for the most anisotropic phase functions, where the conventional cubic spline often develops an oscillating error. The current version of *Phase* calculates up to 2000 non-zero Legendre coefficients using the high-order Lobatto quadrature.

The input data are arranged in three files:

- *Configuration file* (config.par) defines wavelengths, the order of MSH, the incidence-view geometry, the file names of input atmospheric and surface properties, and governs printing of the results.
- *Atmospheric Properties file* describes the model of atmosphere and optical properties of aerosols or clouds. The input for aerosol/clouds can be arranged either in generic form or using microphysical properties. In the first case, the user should provide optical thickness ($\Delta\tau$), single scattering albedo (ω), and scattering function of aerosols/clouds for each atmospheric layer with Legendre expansion computed by program *Phase*. In the second case, the aerosols are assumed to be polydisperse spherical or spheroidal particles with bi-modal lognormal size distribution. The user input includes size distribution parameters and refractive index. The aerosol properties in this case are assumed to be constant with altitude but aerosol concentration may vary according to the specified vertical profile. This model is quite realistic and well suited for the remote sensing studies of atmospheric aerosol and surface reflectance. The single scattering albedo and phase function of aerosol are computed internally with the fast kernel model (Dubovik et al., 2006), followed by automatic Legendre expansion.
- *Surface Properties file* describes the model and parameters of surface reflectance. The details of parameters and input format are documented in 'SHARM Manual'.

Codes Phase, SHARM and SHARM-3D, described in section 6.6, share a common library of files and are written in C language with C++ features.

The Rayleigh scattering is computed automatically according to specified surface pressure and a vertical profile of pressure and temperature, which can be se-

lected from six standard atmospheric profiles (Kneizys et al., 1996) (Tropical, Midlatitude Summer, Midlatitude Winter, SubArctic Summer, SubArctic Winter, 1976 US Standard). The integral over altitude is evaluated with the gaussian quadrature. The Rayleigh scattering cross-section is calculated with the algorithm of Bodhaine et al. (1999), which has a uniformly high accuracy from UV to the shortwave IR.

Three broadly recognized BRF models of land surface reflectance are built in code SHARM: Rahman et al. (1993) (RPV) model, modified RPV (MRPV) (Martonchik et al., 1998) model, and a Li Sparse – Ross Thick reciprocal (LSRT) (Lucht et al., 2000) model. The LSRT and MRPV models are used in the operational land reflectance algorithms of MODIS (Schaaf et al., 2002) and MISR (Martonchik et al., 1998), respectively. All of these models are described by three parameters. They are also reciprocal and rotationally invariant, i.e. they depend only on relative azimuth.

The ocean surface reflectance can be modeled either with the azimuthally-independent model of Nakajima and Tanaka (1983) (NT), or with the Cox and Munk (1954) model with Grams–Charlier expansion (CM). Both models include bi-directional wave-shadowing factor of Nakajima and Tanaka (1983). The NT model depends only on the wind speed, whereas the CM model additionally depends on the wind direction. The details of surface BRF models are given in 'SHARM Manual'.

6.2.1 Accuracy, convergence and speed of SHARM

Code SHARM was extensively validated against code DISORT (Lyapustin, 2002) and was tested experimentally, e.g., over the dark ocean (Gatebe et al., 2005). The most important parameter that controls the accuracy of the solution is the order of MSH (parameter $nb = L$) specified by the user. The single-scattered radiance in SHARM is calculated using an exact formula and does not depend on nb. The solution for the multiple scattering converges to the true one at the increase of nb; however, the computing time also grows approximately as nb^3.

Generally, higher nb is required for more asymmetric phase function and for high zenith angles in order to achieve a given accuracy. For typical continental/marine aerosols, 0.2–0.3% accuracy of radiance is achieved at $nb = 24 - 36$ for view/solar zenith angles up to 75–80°. The value of $nb = 128$ ensures the accuracy of ≈0.02%. Flux calculations require lower nb. For example, fluxes at $nb = 12 - 24$ are typically accurate to the fourth significant digit.

For radiance calculations with strongly asymmetric phase functions typical of clouds, we implemented the Delta-M method (Wiscombe, 1977) that achieves accuracy of about 1% at relatively low orders of MSH, $nb = 32 - 64$, except for the aureole region and some transitional area around it (Lyapustin, 2002). Fig. 6.1 illustrates convergence of SHARM solution for a cirrus cloud with and without Delta-M method.

The performance of code SHARM was tested against code DISORT. In the single angle calculations over a Lambertian surface, the speed of SHARM was found to be a factor of 1.5–6 higher than that of DISORT, depending on the number of atmospheric layers and on the selected order of solution (with $nb = N_{streams}$). This difference increases several times when the surface reflects anisotropically (Lyapustin, 2002). Because SHARM solves the radiative transfer problem for all

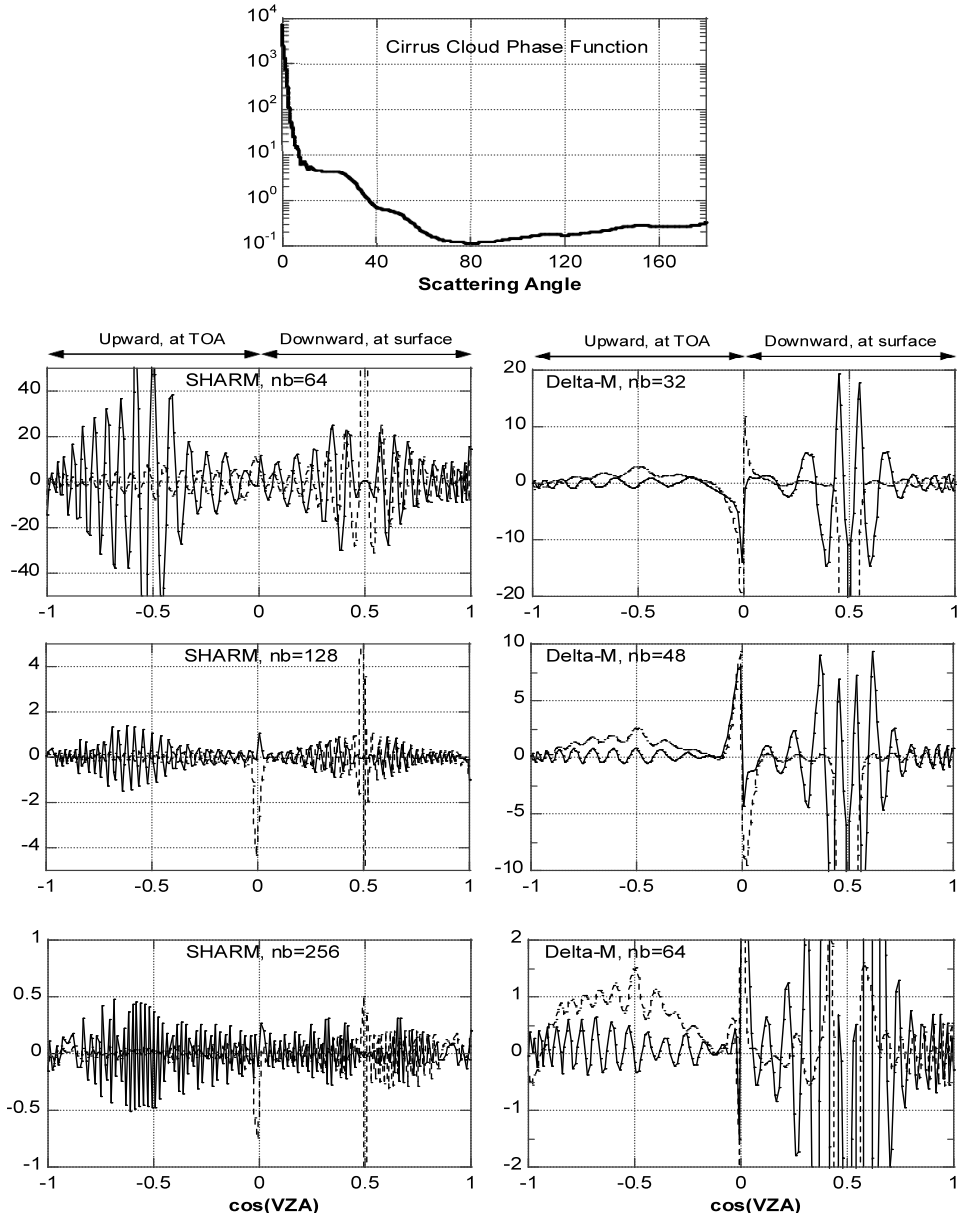


Fig. 6.1. Convergence of SHARM (left) and SHARM with Delta-M method (right) for cirrus cloud (from Lyapustin, 2005). The results are shown as a relative error (%) of SHARM radiance for different orders of MSH calculated with respect to the solution with $nb = 512$. Calculations were performed for $\tau = 0.8$ and $SZA = 60^\circ$. The solid and dashed lines represent the relative azimuth of 0° and 180° respectively. The negative and positive values on x-axis relate to the upward radiance at the top of the atmosphere, and to the diffuse sky radiance incident on surface, respectively. The cirrus cloud phase function at $\lambda = 0.66\,\mu$m corresponds to ice particles with an effective diameter of $10\,\mu$m (http://www.ssec.wisc.edu/∼baum/Cirrus/IceCloudModels.html).

sun-view angles simultaneously, it offers a particular advantage in calculations with complex geometries.

6.3 Green's function method and its applications

In spite of the demonstrated ability of numerical codes for forward simulations, without an analytical solution for non-Lambertian surfaces an efficient atmospheric correction is not possible. This inconvenience transpires in problem of iterative BRF retrieval from multi-angle radiance measurements, e.g. as reported by Lyapustin and Privette (1999). Given atmospheric conditions, the full radiative transfer problem had to be solved anew in each iteration because the lower boundary condition changed.

Below, the Green's function (GF) method is used to derive an analytical solution for the radiance over an anisotropic surface. In the physical sense, the Green's function represents bi-directional atmospheric transmission for the unitary radiance source located at the bottom of the atmosphere. The concept of the Green's function, developed in neutron transport several decades ago (Bell and Glasstone, 1970), offers a powerful approach to solving the radiative transfer problem with complex boundary conditions and internal sources (e.g., Gerstl, 1982; Germogenova, 1986; Sushkevich et al., 1990; Ioltukhovsky, 1999; Landgraf et al., 2002; Qin and Box, 2005; Davis et al., 2009). Below, we use the GF method to derive a semi-analytical solution for TOA radiance as a function of surface BRF (Lyapustin and Knyazikhin, 2001). When used with the linear kernel BRF model (Lucht et al., 2000), this solution yields a formula for TOA radiance which explicitly depends on the BRF model parameters (Lyapustin and Wang, 2005).

6.3.1 Formal solution with the Green's function method

Let us introduce an operator form of notation with differential operator $[\hat{L}_1 = \mu/(\partial/\partial\tau) + 1]$, and integral operators of scattering $[\hat{S} = (\omega/4\pi)\int_\Omega ds'\chi(\tau,\gamma)\dots]$ and of surface reflection $[\hat{R} = \frac{1}{\pi}\int_{\Omega^+} ds'\rho(s',s)\mu'\dots]$, where $s = (\theta,\varphi)$ is a vector of direction. The subscripts/superscripts $(+)$ and $(-)$ will be used to indicate downward $(\mu > 0)$ and upward $(\mu < 0)$ directions, respectively. For example, expression $J_-(\tau_0) = \hat{R}J_+(\tau_0)$ is equivalent to $J(\tau_0;s) = (1/\pi)\int_{\Omega^+} J(\tau_0;s')\rho(s',s)\mu'\,ds'$ at $\mu < 0$.

The Green's function technique allows us to separate the lower boundary condition from calculations of atmospheric functions. First, the TOA radiance is separated into the path radiance $D(\tau,s_0,s)$ and surface-reflected radiance $J(\tau;s_0,s)$:

$$I(\tau;s_0,s) = D(\tau;s_0,s) + J(\tau;s_0,s). \tag{6.21}$$

These components obey the following problems:

$$\hat{L}_1 D = \hat{S}D + S_\lambda \frac{\omega(\tau)}{4}\chi(\tau;\gamma_0)\,\exp(-\tau/\mu_0), \tag{6.22}$$

$$D_+(0) = 0, \quad D_-(\tau_0) = 0; \tag{6.22a}$$

and

$$\hat{L}_1 J = \hat{S} J, \tag{6.23}$$

$$J_+(0) = 0, \quad J_-(\tau_0) = \hat{R} J_+^{(0)}(\tau_0) + \hat{R} J_+(\tau_0). \tag{6.23a}$$

In (6.23a), $J_+^{(0)}(\tau_0)$ is surface irradiance due to directly transmitted sunlight and path radiance,

$$J_+^{(0)}(\tau_0) = \pi S_\lambda \exp(-\tau_0/\mu_0)\delta(s - s_0) + D(\tau_0; s_0, s), \ \mu > 0. \tag{6.24}$$

The term $\hat{R} J_+^{(0)}(\tau_0) = J_-^{(1)}(\tau_0)$ describes the first-order reflection, and the term $\hat{R} J_+(\tau_0)$ describes the multiple photon reflections between the surface and the atmosphere. The solution of (6.23) can be represented as a series in orders of reflection from the surface:

$$J(\tau; s) = \sum_{k \geq 1} J^{(k)}(\tau; s). \tag{6.25}$$

The different orders of reflection ($k \geq 1$) obey the following recurrent system of sub-problems:

$$\hat{L}_1 J^{(k)} = \hat{S} J^{(k)}, \tag{6.26}$$

$$J_+^{(k)}(0) = 0, \quad J_-^{(k)}(\tau_0) = \hat{R} J_+^{(k-1)}(\tau_0). \tag{6.26a}$$

The solution of (6.26) can be analytically expressed with the Green's function $G(\tau; s_1, s)$ of the second kind and boundary values as

$$J^{(k)}(\tau; s) = \int_{\Omega^-} G(\tau; s_1, s) J^{(k)}(\tau_0; s_1) ds_1, \tag{6.27}$$

where $G(\tau; s_1, s)$ satisfies the problem which no longer depends on surface properties:

$$\hat{L}_1 G = \hat{S} G, \tag{6.28}$$

$$G_+(0) = 0, \quad G_-(\tau_0) = \delta(s - s_1), \ \mu < 0. \tag{6.28a}$$

Thus the Green's function describes the solely atmospheric radiative transport and serves to find radiance in an arbitrary direction and altitude in the atmosphere given its angular distribution at the surface (6.27). Separation of the diffuse (continuous) component $G^d(\tau; s_1, s)$ and the direct (discontinuous) component of the Green's function, $G(\tau; s_1, s) = G^d(\tau; s_1, s)$ at $\mu > 0$, and $G(\tau; s_1, s) = \exp(-(\tau_0 - \tau)/|\mu_1|)\delta(s - s_1) + G^d(\tau; s_1, s)$ at $\mu < 0$, gives the transport problem for the diffuse Green's function:

$$\hat{L}_1 G^d = \hat{S} G^d + \frac{\omega(\tau)}{4\pi} \chi(\tau; \gamma_1) \exp(-(\tau_0 - \tau)/|\mu_1|), \tag{6.29}$$

$$G_+^d(0) = 0, \quad G_+^d(\tau_0) = 0, \tag{6.29a}$$

where γ_1 is angle of scattering from the original direction of propagation s_1 into direction s. Problem (6.29) is adjoint to the problem for path radiance (6.22), provided that the source of irradiation is unitary ($\pi S_\lambda = 1$) and located at the

bottom of atmosphere. Therefore, the diffuse Green's function can be computed with the existing numerical codes for the path radiance by reversing the direction of light propagation or, in other words, by setting the atmospheric layers in reverse order, and normalizing the result by πS_λ.

Let us now introduce an integral operator $\hat{\Gamma}_\tau(s)$ corresponding to integral transformation (6.27), $\hat{\Gamma}_\tau(s) = \int_{\Omega^-} ds_1 G(\tau; s_1, s) \ldots$. According to accepted notation, operators $\hat{\Gamma}^+_{\tau_0}$ and $\hat{\Gamma}^-_0$ describe the bidirectional atmospheric transmittance in backward ($\mu > 0$) and forward ($\mu < 0$) directions, respectively, for the source illuminating the atmosphere from below ($\mu_1 < 0$). Using these operators, one can establish the recurrent relation between successive orders of upwelling radiance at the surface level:

$$J^{(k)}_-(\tau_0) = \hat{R}\hat{\Gamma}^+_{\tau_0} J^{(k-1)}_-(\tau_0). \tag{6.30}$$

The total surface-reflected radiance is the sum of all orders of reflection:

$$J_-(\tau_0) = \sum_{k\geq 1} J^{(k)}_-(\tau_0) = \left\{ \sum_{k\geq 0} (\hat{R}\hat{\Gamma}^+_{\tau_0})^k \right\} \hat{R} J^{(0)}_+(\tau_0) = (\hat{I} - \hat{R}\hat{\Gamma}^+_{\tau_0})^{-1} \hat{R} J^{(0)}_+(\tau_0), \tag{6.31}$$

where $\hat{I}$ is a unitary operator. Now, the TOA radiance can be written as

$$I(\tau = 0; s_0, s) = D(0; s_0, s) + \hat{\Gamma}^-_0 (\hat{I} - \hat{R}\hat{\Gamma}^+_{\tau_0})^{-1} \hat{R} J^{(0)}_+(\tau_0), \ \mu < 0. \tag{6.32}$$

Equation (6.32) (Lyapustin and Knyazikhin, 2001) generalizes the well-known formula of Chandrasekhar (1960), originally derived for a Lambertian surface. The inverse operator $(\hat{I} - \hat{R}\hat{\Gamma}^+_{\tau_0,s})^{-1}$ describes multiple reflections of photons between the surface and the atmosphere. When surface is Lambertian, (6.32) transforms into Chandrasekhar's equation with spherical albedo of atmosphere $c_0 = 1/\pi \int_{\Omega^+} \mu' \, ds' \int_{\Omega^-} G^d(\tau_0; s_1, s') \, ds_1$ and upward atmospheric transmittance

$$T(\mu) = \int_{\Omega^-} G(0; s_1, s) \, ds_1 = \exp(-\tau_0/|\mu|) + \int_{\Omega^-} G^d(0; s_1, s) \, ds_1.$$

Let us re-write (6.32) for the TOA radiance in the integral form:

$$I(s_0, s) = D(s_0, s) + \exp(-\tau_0/|\mu|) J(\tau_0; s) + \int_{\Omega^-} G^d(0; s_1, s) J(\tau_0; s_1) ds_1, \quad \mu < 0. \tag{6.33}$$

The total surface-reflected signal is represented by sum (6.31), with the first-order radiance,

$$J^{(1)}(\tau_0; s) = S_\lambda \mu_0 \, \exp(-\tau_0/\mu_0) \rho(s_0, s) + \frac{1}{\pi} \int_{\Omega^+} D(\tau_0; s_0, s') \rho(s', s) \mu' ds'. \tag{6.34}$$

The higher orders of reflection can be computed using (6.30):

$$J^{(k)}(\tau_0; s) = \frac{1}{\pi} \int_{\Omega^-} H(\tau_0; s_1, s) J^{(k-1)}(\tau_0; s_1) \, ds_1, \tag{6.35}$$

where supplementary function H is given by

$$H(\tau_0; s_1, s) = \int_{\Omega^+} G^d(\tau_0; s_1, s')\mu' \rho(s', s)\, ds'. \tag{6.36}$$

The Green's function technique is closely related to the adding-doubling method, e.g. (Hansen and Hovenier, 1971; Twomey, 1985), resulting in similarity of the derived operator equations and the corresponding matrix equations of the adding-doubling method.

6.3.2 Practical considerations

The derived analytical formulas (6.33)–(6.36) can be used to compute atmospheric radiance with accuracy depending on the order of quadrature of integration as well as on the number of multiple reflection terms retained. Computation of successive orders of reflection using (6.35) is impractical, especially over bright surfaces when large number of terms may be required. In practice, significant acceleration can be achieved by using a maximum eigenvalue parameterization. One of the theorems of the linear operator analysis (Vladimirov, 1963; Riesz, Sz.-Nagy, 1990) states that for a continuous linear operator T defined for positive functions u, the following relation holds, starting from some $k > 0$:

$$T^k u(s)/u(s) \cong \eta^k, \tag{6.37}$$

where η is a maximal eigenvalue of operator T. In our case, the lower boundary condition (6.23a) can be represented as a Fredholm equation of the second kind with operator $T = \hat{R}\hat{\Gamma}^+_{\tau_0}$ and $u = J(\tau_0; s)$. Therefore, relationship (6.37) can be used to evaluate the term $k{+}1$ via term k:

$$J^{(k+1)}_-(\tau_0) = (\hat{R}\hat{\Gamma}^+_{\tau_0})J^{(k)}_-(\tau_0) \cong \eta J^{(k)}_-(\tau_0), \tag{6.38}$$

where η is the maximum eigenvalue of operator $\hat{R}\hat{\Gamma}^+_{\tau_0}$, i.e., it is a function of surface reflectance and atmospheric parameters only and does not depend on the view-illumination geometry.

Fig. 6.2 illustrates the total radiance and the first five orders of the surface-reflected radiance computed for grasses in the near-IR range of spectrum ($0.75\,\mu$m). At solar zenith angle of 60° and an atmospheric optical thickness of 0.53, the albedo of grasses is 0.48, so nearly half of the incident solar energy is reflected back into the atmosphere in each instance of reflection. The nonlinear reflection terms in Fig. 6.2 are scaled to show a close similarity in angular shape, starting from the second order of reflection. We found that relation (6.38) holds at $k \geq 2$ to an accuracy of several tenths of a percent for common land surface types, including vegetation, soil, sand and snow. It should be mentioned, however, that the convergence slows down for a ruffled water surface because of a much higher anisotropy of reflectance. In this case, as many as 4–5 orders of reflection may be needed before a constant ratio can be assumed.

This result allows us to introduce an accurate parameterization:

$$J(\tau_0; s) = J^{(1)}(\tau_0; s) + \frac{J^{(2)}(\tau_0; s)}{1 - \eta}, \tag{6.39}$$

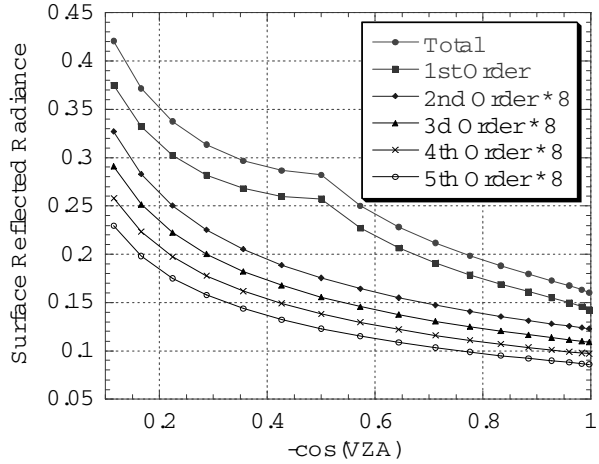


Fig. 6.2. Different orders of the surface-reflected radiance as function of view zenith angle (VZA) at $SZA = 60^\circ$ and $\varphi = 180^\circ$ (from Lyapustin and Knyazikhin, 2001). Calculations were done in the near-IR for continental aerosols with optical thickness 0.5 with surface BRF of grasses. The higher orders of reflection (2–5) are multiplied by a factor of 8 to fit the same scale.

which shows that the surface-reflected radiance over land can be found from the first three orders of reflection, in which the third order is used only to evaluate parameter η. Our analysis for realistic land-cover types showed that with an uncertainty of up to a factor of 2, parameter η can be evaluated in the Lambertian approximation, $\eta \approx q(\mu_0)c_0$, where q is surface albedo. This approximation becomes more accurate as albedo or anisotropy of the BRF decreases.

Equation (6.39) can be further simplified to avoid computing the second-order term. Let us use operator notation and separate incident radiance into the direct beam $I_+^\delta = \pi S_\lambda \exp(-\tau_0/\mu_0)\delta(s - s_0)$ and path radiance:

$$J(\tau_0; s) = \hat{R}[I_+^\delta + D_+(\tau_0)] + \frac{\hat{R}\hat{\Gamma}_{\tau_0}^+\hat{R}[I_+^\delta + D_+(\tau_0)]}{1 - \eta}. \tag{6.40}$$

The reflected diffuse radiance $\hat{R}D_+(\tau_0)$ in the last term is a smooth function in angles, which can be parameterized by the maximum eigenvalue method, $\hat{R}\hat{\Gamma}_{\tau_0}^+\hat{R}D_+(\tau_0) \approx \eta\hat{R}D_+(\tau_0)$. Next, the integration order can be reduced in the term:

$$\begin{aligned}\hat{R}\hat{\Gamma}_{\tau_0}^+\hat{R}I_+^\delta &= S_\lambda\mu_0 \exp(-\tau_0/\mu_0)\hat{R}\hat{\Gamma}_{\tau_0}^+\rho(s_0, s) \\ &= S_\lambda\mu_0 \exp(-\tau_0/\mu_0)\frac{1}{\pi}\int_{\Omega^+}\rho(s', s)\left\{\int_{\Omega^-} G^d(\tau_0; s_1, s')\rho(s_0, s_1)\, ds_1\right\}\mu'\, ds'.\end{aligned} \tag{6.41}$$

This can be done approximately by taking the average value of the Green's function, multiplied by the cosine of the incidence angle, outside of both integral signs:

$$\hat{R}\hat{I}^{+}_{\tau_0}\hat{R}I^{\delta}_{+} \approx S_\lambda\mu_0 \exp(-\tau_0/\mu_0)\left[\frac{1}{\pi}\int_{\Omega^+}\mu'\,ds'\int_{\Omega^-}G^d(\tau_0;s_1,s')\,ds_1\Big/ \left(\int_{\Omega^+}ds'\int_{\Omega^-}ds_1\right)\right]\times\int_{\Omega^+}\rho(s',s)\,ds'\int_{\Omega^-}\rho(s_0,s_1)ds_1$$
$$= S_\lambda\mu_0 \exp(-\tau_0/\mu_0)c_0\rho_1(\mu)\rho_2(\mu_0), \tag{6.42}$$

where

$$\rho_1(\mu) = \int_{\Omega^+}\rho(s',s)\,ds'\Big/\int_{\Omega^+}ds' = \frac{1}{2\pi}\int_{\Omega^+}\rho(s',s)\,ds',$$
$$\rho_2(\mu_0) = \frac{1}{2\pi}\int_{\Omega^-}\rho(s_0,s)\,ds. \tag{6.43}$$

Using assessment $\eta \approx q(\theta_0)c_0$, we arrive at the final formula for the surface-reflected radiance:

$$J(\tau_0;s) \cong S_\lambda\mu_0 \exp(-\tau_0/\mu_0)\big\{\rho(s_0,s)+\alpha c_0\rho_1(\mu)\rho_2(\mu_0)\big\} + \frac{\alpha}{\pi}\int_{\Omega^+}D(\tau_0;s_0,s')\rho(s',s)\mu'\,ds', \tag{6.44}$$

where $\alpha = (1-qc_0)^{-1}$ is a multiple reflection factor. Recalling from (6.33), the total TOA radiance is given by:

$$I(s_0,s) = D(s_0,s) + \exp(-\tau_0/|\mu|)J(\tau_0;s) + \int_{\Omega^-}G^d(0;s_1,s)J(\tau_0;s_1)\,ds_1, \quad \mu<0. \tag{6.45}$$

The surface albedo is defined as a ratio of reflected and incident radiative fluxes at the surface:

$$q(\mu_0) = F^{Up}(\mu_0)/F^{Down}(\mu_0), \tag{6.46}$$

$$F^{Down}(\mu_0) = \pi S_\lambda\mu_0\, e^{-\tau/\mu_0} + \int_{\Omega^+}D_s(s_0,s')\mu'\,ds'$$
$$= F_s^{Dir}(\mu_0) + F_s^{Dif}(\mu_0), \tag{6.44a}$$

$$F^{Up}(\mu_0) = \pi S_\lambda\mu_0\, e^{-\tau/\mu_0}q_2(\mu_0) + \int_{\Omega^+}\mu' q_2(\mu')D_s(s_0,s')\,ds',$$
$$q_2(\mu_0) = \frac{1}{\pi}\int_{\Omega^-}\rho(s_0,s)\mu\,ds. \tag{6.44b}$$

Parametric model (6.41)–(6.44) has a high accuracy, usually within 0.1% to 0.5% for natural landcovers in the bright near-IR region, and better in the visible part of spectrum.

6.3.3 Expression for TOA reflectance using LSRT BRF model

Based on the described semi-analytical solution, we can express TOA reflectance as an explicit function of parameters of the BRF model. We will use a semi-empirical

Li Sparse – Ross Thick (LSRT) BRF model (Lucht et al., 2000). This is a linear model, represented as a sum of Lambertian, geometric-optical, and volume scattering components:

$$\rho(\mu_0,\mu,\varphi) = k^L + k^G f_G(\mu_0,\mu,\varphi) + k^V f_V(\mu_0,\mu,\varphi). \tag{6.45}$$

It uses predefined geometric functions (kernels) f_G, f_V to describe different angular shapes. The kernels are independent of the land conditions. The BRF of a pixel is characterized by a combination of three kernel weights, $\vec{K} = \{k^L, k^G, k^V\}^T$. The LSRT model is used in the MODIS BRF/albedo algorithm (Schaaf et al., 2002).

The substitution of equation (6.45) into (6.41)–(6.44) and normalization to the reflectance units ($R = I_\lambda/(\mu_0 S_\lambda)$) gives the following expressions: the surface-reflected signal at the bottom of atmosphere (6.42) becomes

$$\begin{aligned} R_s(\tau_0;\mu_0,\mu,\varphi) \;=\; & e^{-\tau/\mu_0}\big\{k^L + k^G f_G(\mu_0,\mu,\varphi) + k^V f_V(\mu_0,\mu,\varphi) + \alpha c_0 \rho_1(\mu)\rho_2(\mu_0)\big\} \\ & + \alpha\mu_0^{-1}\big\{k^L E_0^d(\mu_0) + k^G D_G^1(\mu_0,\mu,\varphi) + k^V D_V^1(\mu_0,\mu,\varphi)\big\}. \end{aligned} \tag{6.46}$$

The surface signal diffusely transmitted through the atmosphere (the last term of (6.45)) is given by:

$$\begin{aligned} R_s^d(\mu_0,\mu,\varphi) \;=\; & e^{-\tau/\mu_0} \times \Big\{[k^L G^{av}(\mu) + k^G G_G^1(\mu_0,\mu,\varphi) + k^V G_V^1(\mu_0,\mu,\varphi)] \\ & + \alpha c_0 [k^L G^{av}(\mu) + k^G G_G^{11}(\mu) + k^V G_V^{11}(\mu)]\rho_2(\mu_0)\Big\} \\ & + \alpha\mu_0^{-1}\Big\{k^L E_0^d(\mu_0) G^{av}(\mu) + k^G H_G^1(\mu_0,\mu,\varphi) + k^V H_V^1(\mu_0,\mu,\varphi)\Big\}. \end{aligned} \tag{6.47}$$

The surface albedo is written as:

$$q(\mu_0) = E_0^{-1}(\mu_0)\Big\{\mu_0\, e^{-\tau/\mu_0} q_2(\mu_0) + k^L E_0^d(\mu_0) + k^G D_G^3(\mu_0) + k^V D_V^3(\mu_0)\Big\}. \tag{6.48}$$

Different functions of these equations represent different integrals of the incident path radiance (D_s) and atmospheric Green's function (G) with the BRF kernels which are computed using numerical integration. Below, we only give the integral expressions:

$$\rho_1(\mu) = k^L + k^G f_G^1(\mu) + k^V f_V^1(\mu), \tag{6.49a}$$

$$\rho_2(\mu_0) = k^L + k^G f_G^2(\mu_0) + k^V f_V^2(\mu_0), \tag{6.49b}$$

$$q_2(\mu_0) = k^L + k^G f_G^3(\mu_0) + k^V f_V^3(\mu_0), \tag{6.49c}$$

$$D_k^1(\mu_0,\mu,\varphi-\varphi_0) = \frac{1}{\pi}\int_0^1 \mu'\,d\mu' \int_0^{2\pi} d\varphi' D_s(\mu_0,\mu',\varphi'-\varphi_0) f_k(\mu',\mu,\varphi-\varphi'), \tag{6.50a}$$

$$D_k^3(\mu_0) = \frac{1}{\pi}\int_0^{2\pi} d\varphi' \int_0^1 \mu' f_k^3(\mu') D_s(\mu_0,\mu';\varphi')\,d\mu', \tag{6.50b}$$

$$G^{av}(\mu) = \int_{-1}^0 d\mu_1 \int_0^{2\pi} G^d(\mu_1,\mu,\varphi-\varphi_1)\,d\varphi_1, \tag{6.51a}$$

$$G_k^{11}(\mu) = \int_{-1}^{0} f_k^1(\mu_1) d\mu_1 \int_0^{2\pi} G^d(\mu_1, \mu, \varphi - \varphi_1)\, d\varphi_1, \tag{6.51b}$$

$$G_k^1(\mu_0, \mu, \varphi - \varphi_0) = \int_{-1}^{0} d\mu_1 \int_0^{2\pi} G^d(\mu_1, \mu, \varphi - \varphi_1) f_k(\mu_0, \mu_1, \varphi_1 - \varphi_0)\, d\varphi_1, \tag{6.51c}$$

$$H_k^1(\mu_0, \mu, \varphi - \varphi_0) = \int_{-1}^{0} d\mu_1 \int_0^{2\pi} G^d(\mu_1, \mu, \varphi - \varphi_1) D_k^1(\mu_0, \mu_1, \varphi_1 - \varphi_0)\, d\varphi_1. \tag{6.52}$$

The subscript k in the above expressions refers to either geometric-optical (G) or volumetric (V) kernels, and the supplementary functions of the BRF kernels are given by:

$$f_k^1(\mu) = \frac{1}{2\pi} \int_0^1 d\mu' \int_0^{2\pi} f_k(\mu', \mu, \varphi' - \varphi)\, d\varphi', \tag{6.53a}$$

$$f_k^2(\mu_0) = \frac{1}{2\pi} \int_{-1}^{0} d\mu_1 \int_0^{2\pi} f_k(\mu_0, \mu_1, \varphi_1 - \varphi_0)\, d\varphi_1, \tag{6.53b}$$

$$f_k^3(\mu') = \frac{1}{\pi} \int_{-1}^{0} \mu d\mu \int_0^{2\pi} f_k(\mu', \mu, \varphi - \varphi')\, d\varphi. \tag{6.53c}$$

The diffuse and total spectral surface irradiance are calculated from respective fluxes (6.44a) as:

$$E_0^d(\mu_0) = F^{Dif}(\mu_0)/(\pi S_\lambda), \quad E_0(\mu_0) = F^{Down}(\mu_0)/(\pi S_\lambda). \tag{6.54}$$

Let us re-write these equations separating the kernel weights. First, single out small terms proportional to the product $c_0\rho_2(\mu_0)$ into the nonlinear term:

$$R^{nl}(\mu_0, \mu) = \alpha c_0 \rho_2(\mu_0)\, e^{-\tau/\mu_0} \Big\{ e^{-\tau/|\mu|} \rho_1(\mu) + k^L G^{av}(\mu) + k^G G_G^{11}(\mu) + k^V G_V^{11}(\mu) \Big\}. \tag{6.55}$$

Second, collect all remaining multiplicative factors for the kernel weights:

$$F^L(\mu_0, \mu) = (e^{-\tau/\mu_0} + \alpha \mu_0^{-1} E_0^d(\mu_0))(e^{-\tau/|\mu|} + G^{av}(\mu)), \tag{6.56}$$

$$\begin{aligned} F^k(\mu_0, \mu; \varphi) \;=\; & \Big\{ e^{-\tau/\mu_0} f_k(\mu_0, \mu, \varphi) + \alpha \mu_0^{-1} D_k^1(\mu_0, \mu, \varphi) \Big\} e^{-\tau/|\mu|} \\ & + e^{-\tau/\mu_0} G_k^1(\mu_0, \mu, \varphi) + \alpha \mu_0^{-1} H_k^1(\mu_0, \mu, \varphi), \quad k = V,\ G. \end{aligned} \tag{6.57}$$

With these notations, the TOA reflectance becomes:

$$\begin{aligned} R(\mu_0, \mu, \varphi) \;=\; & R^D(\mu_0, \mu, \varphi) + k^L F^L(\mu_0, \mu) + k^G F^G(\mu_0, \mu, \varphi) \\ & + k^V F^V(\mu_0, \mu, \varphi) + R^{nl}(\mu_0, \mu). \end{aligned} \tag{6.58}$$

This equation represents TOA reflectance as an explicit function of the BRF model parameters providing the means for an efficient atmospheric correction. Equation (6.67) is used in the AERONET-based Surface Reflectance Validation Network (ASRVN) to derive spectral surface BRF and albedo from operational MODIS data at AERONET locations (Wang et al., 2009). The developed algorithm uses

AERONET (Holben et al., 1998) aerosol and water vapor data to perform atmospheric correction. The released ASRVN dataset (MODIS product MODASRVN) contains MODIS TERRA and AQUA data since 2000 and 2002, respectively, for $50 \times 50\,\mathrm{km}^2$ area of ∼100 AERONET stations globally. This dataset can be used for different applications, including validation and science analysis.

Developed analytical theory also became a foundation for the Multi-Angle Implementation of Atmospheric Correction algorithm (MAIAC) (Lyapustin and Wang, 2009) which uses time series analysis of MODIS data to simultaneously retrieve surface BRF/albedo and properties of atmospheric aerosol at high 1 km resolution.

6.4 Green's function solution for anisotropic inhomogeneous surface

Land surface is not homogeneous. Spatial variability of surface reflectance gives rise to horizontal radiative fluxes in the atmosphere directed from the bright to the dark surface areas. These three-dimensional (3-D) effects reduce the apparent top-of-the-atmosphere surface contrast by decreasing the radiance over bright pixels and increasing the brightness of the dark pixels (Otterman and Fraser, 1979; Mekler and Kaufman, 1980; Diner and Martonchik, 1985). This blurring effect is systematic and thus becomes important for the remote sensing applications developed for use with either dark or bright targets. Lyapustin (2001) and Lyapustin and Kaufman (2001) showed that 3-D effects may cause a systematic overestimation of the retrieved aerosol optical thickness over land by the dark dense vegetation method and a non-negligible systematic error in the land surface albedo at a broad range of sensor resolutions in many climatically important regions with medium-to-high contrast of surface.

One of the first numerical solutions for problem with nonhomogeneous non-Lambertian boundary was obtained by Diner and Martonchik (1984a,b). A number of studies were done with Monte Carlo techniques (Pearce, 1977; Marchuk et al., 1980) using Lambertian surface model. One of the most versatile codes for 3-D analysis, the spherical harmonic discrete ordinates method (SHDOM), was developed by Evans (1998). SHDOM handles the complete 3-D problem with the lateral variability of both surface and atmospheric optical properties.

An exact semi-analytical solution obtained with the Green's function method along with the method of successive surface interactions (Lyapustin and Knyazikhin, 2002) is described next. Following methodology used in the 1-D case (section 3), parameterizations are developed which allow a fast numerical implementation with high accuracy suitable for the remote sensing applications.

6.4.1 Operator solution of the 3-D radiative transfer problem

The diffuse solar radiation in the horizontally homogeneous atmosphere with non-uniform non-Lambertian surface is a solution of the following 3-D boundary-value problem:

$$[(\vec{\nu}\vec{\nabla}) + \alpha(z)]I(z;r;s) = \frac{\sigma(z)}{4\pi}\int_{\Omega}\chi(z,\gamma)I(z;r;s')\,ds' + \frac{S_\lambda}{4}\sigma(z)\chi(z;\gamma_0)\exp(-\tau(z)/\mu_0), \tag{6.59}$$

$$I(0;r;s) = 0, \quad \mu > 0, \tag{6.59a}$$

$$I(H;r;s) = S_\lambda\mu_0\,\exp(-\tau_0/\mu_0)\rho(r;s_0,s) + \frac{1}{\pi}\int_{\Omega^+} I(H;r;s')\rho(r;s',s)\mu'\,ds',\ \mu < 0. \tag{6.59b}$$

Here, $\sigma(z)$ and $\alpha(z)$ are scattering and extinction coefficients of the atmosphere at altitude z, respectively, $\vec{\nu} = (\sqrt{1-\mu^2}\,\cos\varphi, \sqrt{1-\mu^2}\,\sin\varphi, \mu)$ is vector of direction, $\vec{\nabla} = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ is vector of gradient, and $r = (x, y)$ is a horizontal coordinate. As before, we separate the path radiance $D(z, s_0, s)$ and surface-reflected radiance $J(r; z; s)$ in the total signal, and will use the operator notations with the 3-D differential operator $[\hat{L}_3 = (\vec{\nu}\vec{\nabla}) + \alpha(z)]$ and integral operators of atmospheric scattering

$$\left[\hat{S} = \frac{\sigma(z)}{4\pi}\int_{\Omega} ds'\chi(z,\gamma)\ldots\right]$$

and of surface reflection

$$\left[R_r = \frac{1}{\pi}\int_{\Omega^+} ds'\mu'\rho(r;s',s)\ldots\right].$$

With these notations, the problem for the surface-reflected radiance becomes:

$$\hat{L}_3 J(z;r;s) = \hat{S}J(z;r;s), \tag{6.60}$$

$$J_+(0;r) = 0, \quad J_-(H;r) = R_r J_+^{(0)}(H) + R_r J_+(H;r). \tag{6.60a}$$

Expansion of solution in a series of surface interactions:

$$J(z;r;s) = \sum_{k\geq 1} J^{(k)}(z;r;s), \tag{6.61}$$

generates a set of recursive problems for different orders of reflection $J^{(k)}$:

$$\hat{L}_3 J^{(k)}(z;r;s) = \hat{S}J^{(k)}(z;r;s), \tag{6.62}$$

$$J_+^{(k)}(0;r) = 0; \quad J_-^{(k)}(H;r) = R_r J_+^{(k-1)}(H;r). \tag{6.62a}$$

As before, the solution is expressed via 3-D surface Green's function $G_3(z; r - r'; s_1, s)$ and the boundary values of radiance $J^{(k)}(H; r; s_1)$:

$$J^{(k)}(z;r;s) = \int_{-\infty}^{+\infty} dr' \int_{\Omega^-} G_3(z; r - r'; s_1, s)J^{(k)}(H; r'; s_1)\,ds_1. \tag{6.63}$$

The substitution of (6.63) into problem (6.62) shows that the Green's function thus introduced does not depend on the reflective properties of the surface and satisfies a classical searchlight problem originally introduced by Chandrasekhar (1958):

$$\hat{L}_3 G_3 = \hat{S} G_3, \tag{6.64}$$

$$G_{+3}(0; r - r') = 0, \quad G_{-3}(H; r - r') = \delta(r - r')\delta(s - s_1). \tag{6.64a}$$

The lower boundary condition shows that function $G_3(z; r - r'; s_1, s)$ is an atmospheric response to the unitary mono-directional localized perturbation on its lower boundary. For this reason, function G_3 is also called a point-spread function (PSF).

Let us introduce an operator corresponding to integral transformation (6.63)

$$\hat{\Gamma}^3_{z;r}(s) = \int_{-\infty}^{+\infty} dr' \int_{\Omega^-} ds_1 G_3(z; r - r'; s_1, s) \dots \ . \tag{6.65}$$

The operators $\hat{\Gamma}^{3+}_{H;r} \equiv \hat{\Gamma}^3_{H;r}(\mu > 0, \varphi)$ and $\hat{\Gamma}^{3-}_{0;r} \equiv \hat{\Gamma}^3_{z=0;r}(\mu < 0, \varphi)$ represent atmospheric transmittance for radiance in the backward and forward directions, respectively, upon illumination of the atmosphere from below. Now we can establish the relation between successive orders of interaction of light with inhomogeneous surface. At the surface level, the downward radiance of the reflection order (k), $J^{(k)}_+(H; r)$, is a result of atmospheric backscattering of the upward radiance of the same order, $J^{(k)}_+(H; r) = \hat{\Gamma}^{3+}_{H;r} J^{(k)}_-(H; r)$. The upward radiance of the order $(k+1)$ can be expressed via its precursor of the order (k) as:

$$J^{(k+1)}_-(H; r) = R_r J^{(k)}_+(H; r) = R_r \hat{\Gamma}^{3+}_{H;r} J^{(k)}_-(H; r). \tag{6.66}$$

Based on this result, the total upward surface-reflected radiance in series (6.61) becomes

$$\begin{aligned} J_-(H; r) = \sum_{k \geq 1} J^{(k)}_-(H; r) \ &= \ \sum_{k \geq 0} (R_r \hat{\Gamma}^{3+}_{H;r})^k R_r J^{(0)}_+(H) \\ &= \ (\hat{I} - R_r \hat{\Gamma}^{3+}_{H;r})^{-1} R_r J^{(0)}_+(H). \end{aligned} \tag{6.67}$$

This formula gives an exact formal solution in the operator form for the reflected radiance at the surface level. The relationships between successive orders of interaction $k \to k + 1$ are described by the operator $R_r \hat{\Gamma}^{3+}_{H;r}$, which represents the following integral transformation:

$$\begin{aligned} J^{(k+1)}(H; r; s) \ &= \ \frac{1}{\pi} \int_{\Omega^+} ds' \mu' \rho(r; s', s) \\ &\times \ \int_{-\infty}^{+\infty} dr' \int_{\Omega^-} G_3(H; r - r'; s_1, s') J^{(k)}(H; r'; s_1)\, ds_1. \end{aligned} \tag{6.68}$$

At an arbitrary altitude z in the atmosphere, the surface-reflected radiance in the upward direction can be found from (6.63):

$$J_-(z; r - r_s) = \hat{\Gamma}^{3-}_{z;r-r_s} (\hat{I} - R_r \hat{\Gamma}^{3+}_{H,r})^{-1} R_r J^{(0)}_+(H), \tag{6.69}$$

where $r_s = \{(H - z)\,|tg\theta|\,\cos\varphi; (H - z)\,|tg\theta|\,\sin\varphi\}$ is a shift in the horizontal coordinate at slant observations.

6.4.2 Linearized solution

An exact solution (6.69) accounts for all orders of interaction of light with spatially variable surface BRF. In further analysis, it is convenient to separate the mean value $[\bar{f}]$ and the spatial variation $[\tilde{f}(r)]$ in surface BRF and in radiance. We will further analyze (6.69) in the linear approximation in BRF's spatial variation. It will allow us to develop a fast parametric solution and to use an earlier-derived parametric formula for the mean component of radiance (section 6.3). The linearized solution has a sufficiently high accuracy for practical purposes. The non-linear contribution $\tilde{J}_{nl}(z;r;s)$ is typically small. For example, when the surface is Lambertian, the non-linear contribution is bounded (Sushkevich et al., 1990):

$$\tilde{J}_{nl}(z;r;s)/\tilde{J}(z;r;s) \le \frac{q^{\max}c_0}{1-\bar{q}c_0}, \tag{6.70}$$

where q^{max} is a maximal albedo value in the image. In clear-sky conditions, $\tilde{J}_{nl}(z;r;s)$ does not exceed several percent of the radiance variation and is even smaller with respect to the total signal.

Let us divide the surface reflection operator (R_r) into the mean $(\bar{R})$ and variation $(\tilde{R}_r)$ components, and linearize the surface-reflected radiance (equation (6.67)):

$$\begin{aligned} J_-(H;r) \;=\; & \sum_{k\ge 0}([\bar{R}+\tilde{R}_r]\hat{\Gamma}^{3+}_{H;r})^k[\bar{R}+\tilde{R}_r]J^{(0)}_+(H) \cong \bar{J}_-(H) \\ & + \sum_{k\ge 0}(\bar{R}\hat{\Gamma}^{3+}_{H;r})^k\tilde{R}_r J^{(0)}_+(H) \\ & + \sum_{k\ge 0}\sum_{l=0}^{k}(\bar{R}\hat{\Gamma}^{3+}_{H;r})^{k-l}\tilde{R}_r\hat{\Gamma}^{3+}_{H;r}(\bar{R}\hat{\Gamma}^{3+}_{H;r})^l\bar{R}J^{(0)}_+(H). \end{aligned} \tag{6.71a}$$

In this formula, the first term $\bar{J}_-(H)$ is the mean reflected radiance corresponding to the mean BRF. Let us rewrite the third term as

$$\sum_{k\ge 0}\sum_{l=0}^{k}(\bar{R}\hat{\Gamma}^{3+}_{H;r})^{k-l}\tilde{R}_r(\hat{\Gamma}^{1+}_H\bar{R})^l\hat{\Gamma}^{1+}_H\bar{R}J^{(0)}_+(H). \tag{6.71b}$$

Here, we used the fact that the 3-D operator $\hat{\Gamma}^{3+}_{H;r}$, when applied to the mean signal $\bar{R}J^{(0)}_+(H)$, turns into a 1-D operator $\hat{\Gamma}^{1+}_H$ related to the 1-D Green's function

$$G_1(z;s_1,s) = \int_{-\infty}^{+\infty} G_3(z;r-r';s_1,s)\,dr', \tag{6.72}$$

that obeys the boundary-value problem (6.28). Now, series (6.71b) can be summed analytically as follows. Let us fix index $l = l_0$ $(l_0 = 0, 1, \dots)$, and sum the series in index k from $k = l_0$ to $k = \infty$. When the residual series is summed in index l including the second term of formula (6.71a), it gives the following expression for the variation of surface-reflected radiance:

$$\tilde{J}_-(H;r) \cong \quad (\hat{I}-\bar{R}\hat{\Gamma}^{3+}_{H;r})^{-1}\tilde{R}_r(\hat{I}-\hat{\Gamma}^{1+}_H\bar{R})^{-1}J^{(0)}_+(H). \tag{6.73}$$

This equation shows that in the linear approximation, the variation of the surface-reflected radiance is formed as follows. First comes the mean incident radiance that includes all orders of interaction with the mean surface BRF, $(\hat{I} - \hat{\Gamma}_H^{1+}\bar{R})^{-1}J_+^{(0)}(H)$. The incident radiance is then reflected from the BRF variation, $\tilde{R}_r(\hat{I} - \hat{\Gamma}_H^{1+}\bar{R})^{-1}J_+^{(0)}(H)$. Finally, the variation of reflected signal is enhanced in the process of subsequent multiple interactions with the mean BRF.

Applying a spatial Fourier transform, we can write an expression for the spatial Fourier spectrum of radiance variation at altitude z:

$$\tilde{J}_-(z;p) = F[\tilde{J}_-(z;r)] = \hat{\Gamma}_{z;p}^{3-}(\hat{I} - \bar{R}\hat{\Gamma}_{H;p}^{3+})^{-1}\tilde{R}_p(\hat{I} - \hat{\Gamma}_H^{1+}\bar{R})^{-1}J_+^{(0)}(H). \quad (6.74)$$

Here, the Fourier transform is defined for spatial frequency $p = (p_x, p_y)$ as $F[y(r)] = y(p) = \int_{-\infty}^{+\infty} y(r)\,\exp(ipr)\,dr$, and operator $\hat{\Gamma}_{z;p}^{3-}$ is defined as

$$\hat{\Gamma}_{z;p}^{3-} = \int_{\Omega^-} ds_1 G_3(z;p;s_1,s)\ldots,\ \mu < 0 \quad (6.75)$$

The Fourier transform of PSF, $G_{3p} \equiv G_3(z;p;s_1,s)$, is called an optical transfer function (OTF) of the atmosphere. The boundary-value problem for OTF is obtained by applying a Fourier transform to the problem for PSF (6.64):

$$\hat{L}_{3p}G_{3p} = \hat{S}G_{3p}, \quad (6.76)$$

$$G_{+3p}(0) = 0; \quad G_{-3p}(H) = \delta(s - s_1), \quad (6.76a)$$

where $\hat{L}_{3p} = \mu\partial/\partial z - i\sqrt{1-\mu^2} \times \{p_x\cos\varphi + p_y\sin\varphi\} + \alpha(z)$. In the downward direction, OTF is a purely diffuse function G_{3p}^d, and it has both direct and diffuse components in the upward direction:

$$\begin{aligned} &G_3(z;p;s_1,s) \\ &= \exp(ipr_s)\begin{cases} G_3^d(z;p;s_1,s),\ \mu > 0 \\ \exp(-[\tau_0 - \tau(z)]/|\mu_1|)\delta(s-s_1) + G_3^d(z;p;s_1,s),\ \mu < 0 \end{cases}. \end{aligned} \quad (6.77)$$

Since operator $\hat{L}_{3p}$ is complex, the diffuse component of OTF is also a complex function and can be expressed via its amplitude A and phase Φ, $G_3^d(z;p;s_1,s) = A(z;p;s_1,s)\,\exp(i\Phi(z;p;s_1,s))$.

The total radiance at altitude z can be obtained using inverse Fourier transform:

$$I(z;r;s) = D(z;s_0,s) + \bar{J}(z;s_0,s) + F^{-1}[\tilde{J}(z;p;s)]. \quad (6.78)$$

Let us now consider a linearized solution in application to a Lambertian surface, which has been rather well studied both in theory (Sushkevich et al., 1990; Zege et al., 1991) and in numerical experiments (Lyapustin, 2001; Lyapustin and Muldashev, 2001). This exercise will serve a twofold purpose. On the one hand, we will show that the Lambertian approximation is a particular case of a general solution (6.74). On the other hand, based on the Lambertian solution, we will introduce important parameterizations that will allow us to significantly accelerate calculations in the general case of anisotropic surface reflectance.

6.4.3 Lambertian approximation

In the case of isotropic surface reflectance, formula (6.74) transforms as follows.

(1) The variation of the reflected radiance becomes

$$\begin{aligned}\tilde{R}_p(\hat{I}-\hat{\Gamma}_H^{1+}\bar{R})^{-1}J_+^{(0)}(H) &= \tilde{R}_p\sum_{k\geq 0}(\hat{\Gamma}_H^{1+}\bar{R})^k J_+^{(0)}(H)\\ &= \tilde{q}(p)\sum_{k\geq 0}(c_0\bar{q})^k E_0(\mu_0)\\ &= \tilde{q}(p)\alpha E_0(\mu_0), \qquad (6.79a)\end{aligned}$$

where surface irradiance

$$E_0(\mu_0) = S_\lambda\mu_0\exp(-\frac{\tau_0}{\mu_0}) + \frac{1}{\pi}\int_{\Omega^+} D(H;s_0,s)\mu\, ds,$$

and $\alpha = (1-\bar{q}c_0)^{-1}$.

(2) Operator $\bar{R}\hat{\Gamma}_{H;p}^{3+}$, coupling atmospheric backscattering and reflection from the mean BRF, modifies variation of surface-reflected radiance as follows:

$$\begin{aligned}&(\bar{R}\hat{\Gamma}_{H;p}^{3+})\tilde{R}_p(\hat{I}-\hat{\Gamma}_H^{1+}\bar{R})^{-1}J_+^{(0)}(H)\\ &\quad= \frac{\bar{q}}{\pi}\int_{\Omega^+}\int_{\Omega^-}\mu' G_3^d(H;p;s_1,s')ds_1ds'\times\tilde{q}(p)\alpha E_0(\mu_0)\\ &\quad= \{\bar{q}c(p)\}\tilde{q}(p)\alpha E_0(\mu_0), \qquad (6.79b)\end{aligned}$$

where

$$c(p) = \frac{1}{\pi}\int_{\Omega^+}\mu' ds'\int_{\Omega^-} G_3^d(H;p;s_1,s')\, ds_1$$

is a spherical albedo of atmosphere at spatial frequency p. At $p\to 0$, it turns into a 1-D spherical albedo of atmosphere, $c_0 \equiv c(0)$.

Equation (6.79b) allows us to express the linear variation of surface-reflected radiance as:

$$\begin{aligned}\tilde{J}(H;r;s) &= \sum_{k\geq 0}(\bar{R}\hat{\Gamma}_{H;r}^{3+})^k\tilde{R}_p(\hat{I}-\hat{\Gamma}_H^{1+}\bar{R})^{-1}J_+^{(0)}(H)\\ &= \sum_{k\geq 0}\{\bar{q}c(p)\}^k\times\tilde{q}(p)\alpha E_0(\mu_0) = \tilde{q}_m(p)\alpha E_0(\mu_0), \qquad (6.79c)\end{aligned}$$

where $\tilde{q}_m(p) = [\tilde{q}(p)]/[1-\bar{q}c(p)]$ is the Fourier transform of albedo variation corrected for multiple interactions between the atmosphere and the mean surface albedo.

(3) Since the angular distribution of reflected radiance $\tilde{J}(H;r;s)$ is isotropic, operator $\hat{\Gamma}_{z;p}^{3-}$ becomes a scalar function

$$\begin{aligned}&\hat{\Gamma}_{z;p}^{3-}(I-\bar{R}\hat{\Gamma}_{H;p}^{3+})^{-1}\tilde{R}_p(\hat{I}-\hat{\Gamma}_H^{1+}\bar{R})^{-1}J_+^{(0)}(H)\\ &\quad= \tilde{q}_m(p)\alpha E_0(\mu_0)\int_{\Omega^-} G_3(z;p;s_1,s)ds_1\\ &\quad= \tilde{q}_m(p)\alpha E_0(\mu_0)\Psi(z;p;s). \qquad (6.79d)\end{aligned}$$

Here,

$$\Psi(z;p;s) = \int_{\Omega^-} G_3(z;p;s_1,s)\,ds_1 = \exp(ipr_s)\begin{cases} \int_{\Omega^-} G_3^d(z;p;s_1,s)ds_1, \mu>0 \\ \exp(-[\tau_0-\tau(z)]/|\mu| + \int_{\Omega^-} G_3^d(z;p;s_1,s)ds_1, \mu<0 \end{cases}, \tag{6.79e}$$

is atmospheric OTF for the problem with Lambertian surface. From (6.79e) and (6.76) it follows that function $\Psi_p \equiv \Psi(z;p;s)$ obeys the problem with isotropic unitary source on the lower boundary:

$$\hat{L}_{3p}\Psi_p = \hat{S}\Psi_p, \tag{6.80}$$

$$\Psi_p^+(0) = 0, \quad \Psi_p^-(H) = 1. \tag{6.80a}$$

(4) Finally, the inverse Fourier transform of Eq. (6.79d) gives the radiance variation in the upward direction,

$$\tilde{I}(z;r;s) = \alpha E_0(\mu_0)\Big\{\tilde{q}_m(r-r_s)\exp(-[\tau_0-\tau(z)]/|\mu|) + \frac{1}{(2\pi)^2}\int_{-\infty}^{+\infty}\tilde{q}_m(p)A(z;p;s)\,\exp(-i[p(r-r_s)-\Phi(z;p;s)])\,dp\Big\}, \tag{6.81}$$

where A and Φ are the amplitude and phase of the diffuse component of OTF $\Psi(z;p;s)$.

The boundary-value problem (6.80) defining OTF $\Psi(z;p;s)$, and solution for the radiance variation (6.81) were obtained earlier for the linearized radiative transfer problem with Lambertian lower boundary (Sushkevich et al., 1990). Our independent derivation thus shows that the Lambertian solution is a particular case of a general solution (6.74).

6.4.4 Numerical aspects

Equation (6.74) cannot be used in practice. The evaluation of multiple reflections from surface (inverse operator) and the Fourier transforms are the major computational challenge. Several parameterizations can be introduced that accelerate the solution dramatically while keeping the accuracy sufficiently high for most applications.

First, we will use the maximum eigenvalue method to parameterize multiple reflections in the term

$$\tilde{R}_p(\hat{I}-\hat{\Gamma}_H^{1+}\bar{R})^{-1}J_+^{(0)}(H) = \tilde{R}_pJ_+^{(0)}(H) + \tilde{R}_p\sum_{k\geq 1}(\hat{\Gamma}_H^{1+}\bar{R})^kJ_+^{(0)}(H). \tag{6.82a}$$

Separating the incident radiance $J_+^{(0)}(H)$ into the direct solar beam I_δ^+ and incident path radiance D^+, we can rewrite the last term as (see derivation (6.40)–(6.41))

$$\tilde{R}_p \sum_{k\geq 0} (\hat{\Gamma}_H^{1+}\bar{R})^k \hat{\Gamma}_H^{1+}\bar{R}J_+^{(0)}(H) \cong \frac{\tilde{R}_p\hat{\Gamma}_H^{1+}\bar{R}(I_\delta^+ + D^+)}{1-\bar{\eta}}, \tag{6.82b}$$

where $\bar{\eta}$ is a maximum eigenvalue of operator $\hat{\Gamma}_H^{1+}\bar{R}$. Next, following derivation of section 6.3, we can write:

$$\tilde{R}_p\hat{\Gamma}_H^{1+}\bar{R}I_\delta^+ \approx S_\lambda\mu_0 \exp(-\tau_0/\mu_0)c_0\tilde{\rho}_1(p;\mu)\bar{\rho}_2(\mu_0), \text{ and } \tilde{R}_p\hat{\Gamma}_H^{1+}\bar{R}D^+ \approx \bar{\eta}\tilde{R}_pD^+, \tag{6.82c}$$

where $\rho_1(\mu)$, $\rho_2(\mu_0)$ are given by (6.43). Evaluation of the maximum eigenvalue $\bar{\eta}$ in Lambertian approximation ($\bar{\eta} \approx \bar{q}(\mu_0)c_0$) allows us to rewrite (6.82a) as:

$$\begin{aligned}\tilde{R}_p(\hat{I}-\hat{\Gamma}_H^{1+}\bar{R})^{-1}J_+^{(0)}(H) \;\cong\; & S_\lambda\mu_0 \exp\left(-\frac{\tau_0}{\mu_0}\right)\Big\{\tilde{\rho}(p;s_0,s) + \alpha c_0\tilde{\rho}_1(p;s)\bar{\rho}_2(s_0)\Big\} \\ & + \frac{\alpha}{\pi}\int_{\Omega^+} D(H;s_0,s')\tilde{\rho}(p;s',s)\mu'\,ds', \\ & \alpha = (1-\bar{q}(\mu_0)c_0)^{-1}. \end{aligned} \tag{6.82d}$$

Equation (6.82d) gives the linearized variation of the surface reflected radiance $\tilde{J}(H;p;s)$ without the multiple reflection contribution from the term $(\hat{I}-\bar{R}\hat{\Gamma}_{H;p}^{3+})^{-1}$. Neglecting this term for now, and recalling (6.77)–(6.78), we can express the TOA variation of radiance as:

$$\begin{aligned}\tilde{J}(0;r;s) \;\cong\; & F^{-1}[\exp(ipr_s)\{\tilde{J}(H;p;s)\,\exp(-\tau_0/\,|\mu|) \\ & + \int_{\Omega^-} G_{3p}^d(0;p;s_1,s)\tilde{J}(H;p;s_1)\,ds_1\}]. \end{aligned} \tag{6.83}$$

A quick analysis of this equation shows that the Fourier transformation of the BRF variation is a major computational issue. For example, separate direct and inverse FFT operations are required not only for a given solar angle and view directions, but also for all directions of numerical quadrature necessary for integration over the hemispheres Ω^+, Ω^-, which would make this method completely impractical. However, all of the burden in practice can be reduced to a single direct and inverse Fourier transform.

To demonstrate this, let us consider the second (diffuse) term in formula (6.83). In the physical sense, it represents photons collected from a large elliptic area surrounding the pixel $(r-r_s)$. These photons fall into the line of sight after a single out or several acts of scattering in the atmosphere. In each instance of scattering, they 'lose memory' of the original direction of propagation. The photons also arrive from different surface points, each of which is characterized by its unique bi-directional reflectance and geometry of reflection. Thus, combination of the horizontal and angular averaging due to scattering in the atmosphere produces an effective smoothing of the diffuse transmission of the variation of reflected radiance. Thus, the diffuse term of (6.83) can be modeled in the Lambertian approximation (second term of (6.81)). To this end, we introduce an effective Lambertian surface with albedo distribution defined from the condition of conservation of the specific reflected flux in each surface point:

$$q(\mu_0;r) = F^{\uparrow}(H;r)/\bar{F}^{\downarrow}(H;\mu_0), \tag{6.84a}$$

$$\bar{F}^{\downarrow}(H;\mu_0) = \pi E_0(\mu_0) = \pi S_\lambda \mu_0 \exp(-\tau_0/\mu_0) + \int_{\Omega^+} D(H;s_0,s')\mu'\,ds', \quad (6.84b)$$

$$\begin{aligned} F^{\uparrow}(H;r) \;=\; & \int_{\Omega^-} \mu ds \{ S_\lambda \mu_0 \exp(-\tau_0/\mu_0)\rho(r;s_0,s) \\ & + \frac{1}{\pi}\int_{\Omega^+} \mu'\rho(r;s',s)D(H;s_0,s')\,ds'\}. \end{aligned} \quad (6.84c)$$

In the next step of eliminating multiple Fourier transforms, we also neglect the multiple reflection term $(\hat{I} - \bar{R}\hat{\bar{I}}^{3+}_{H;p})^{-1}$ in the directly transmitted variation of the surface-reflected radiance. Then, the final expression for the TOA radiance becomes:

$$\begin{aligned} I(r-r_s;s_0,s) \;\cong\; & D(0;s_0,s) + \exp(-\tau_0/\,|\mu|)J(H;r;s) + \int_{\Omega^-} G_1^d(0;s_1,s)\bar{J}(H,s_1)\,ds_1 \\ & + \frac{\alpha E_0(\mu_0)}{(2\pi)^2}\int_{-\infty}^{+\infty} \frac{\tilde{q}(\mu_0,p)A(0;p;s)}{1-\bar{q}(\mu_0)c(p)} \exp(-i[p(r-r_s) - \Phi(0;p;s)])\,dp, \end{aligned} \quad (6.85)$$

where surface-reflected radiance is given by

$$\begin{aligned} J(H;r;s) \;\cong\; & S_\lambda \mu_0 \exp(-\tau_0/\mu_0)\Big\{\rho(r;s_0,s) + \alpha c_0 \rho_1(r;\mu)\bar{\rho}_2(\mu_0)\Big\} \\ & + \frac{\alpha}{\pi}\int_{\Omega^+} D(H;s_0,s')\rho(r;s',s)\mu'\,ds'. \end{aligned} \quad (6.86)$$

Equation (6.85) shows that the TOA radiance is represented physically as a sum of path radiance and surface-reflected radiance that is directly and diffusely transmitted through the atmosphere. The diffusely transmitted component is additionally split into the diffuse transmission of the mean value and of variation of the surface-reflected radiance.

Radiance simulations over spatially varying anisotropic surface based on (6.85)–(6.86) require a set of 1-D functions, including path radiance, 1-D Green's function, incident and reflected surface fluxes and spherical albedo of atmosphere, which are computed with code SHARM. It also requires a 3-D optical transfer function of atmosphere, solution for which, by the method of spherical harmonics obtained by Lyapustin and Muldashev (2001), is described next.

6.5 MSH solution for the optical transfer function

Let us consider the boundary-value problem for OTF (6.80). For simplicity, we assume that surface reflectance changes only in the x-direction, which gives a 2-D problem in coordinates $(z, p = p_x)$. Let us re-write (6.80), dividing the equation by extinction coefficient $\alpha(z)$ and defining parameter $p_1 = p/\alpha(z)$:

$$\mu\frac{\partial\psi(\tau;p;s)}{\partial\tau} + [1 - ip_1\sqrt{1-\mu^2}\,\cos\varphi]\psi(\tau;p;s) = \frac{\omega(\tau)}{4\pi}\int_{\Omega}\chi(\tau,\gamma)\psi(\tau;p;s')\,ds' \quad (6.87)$$

$$\psi(0;p;s)=0,\ \mu>0;\quad \psi(\tau_0;p;s)=1,\ \mu<0. \tag{6.87a}$$

Expansion of the scattering function (6.3) and the OTF into azimuthal Fourier series,

$$\psi(\tau;p;s)=(2-\delta_{0,m})\sum_{m=1}^{\infty}\psi^m(\tau;p;\mu)\cos m\varphi, \tag{6.88}$$

provides an equation for the mth azimuthal harmonic ψ^m:

$$\mu\frac{\partial\psi^m}{\partial\tau}+\psi^m-ip_1\sqrt{1-\mu^2}\frac{1+\delta_{0,m}}{2}(\psi^{m-1}+\psi^{m+1})=\frac{\omega}{2}\int_{-1}^{+1}\psi^m\chi^m\,d\mu' \tag{6.89}$$

$$\psi^m(0;p;\mu)=0,\ \mu>0;\quad \psi^m(\tau_0;p;\mu)=\delta_{0,m},\ \mu<0,\ m=0,\ldots,M. \tag{6.89a}$$

Contrary to the 1-D problem for mean intensity, in this case equations at different m are chained, so system (6.89) can only be solved simultaneously for all $m=0,\ldots,M$.

The solution for the OTF is a complex function, $\psi^m=\psi^m_{Re}+i\psi^m_{Im}$. From the fact that there are no internal sources in (6.89), and from the form of boundary conditions, it follows that $\psi^{2k+1}_{Re}=0$, and $\psi^{2k}_{Im}=0$, $k=0,\ldots,M/2$. So we can define a new real function Φ according to the rule: $\Phi^{2k}=\psi^{2k}_{Re}$, and $\Phi^{2k+1}=\psi^{2k+1}_{Im}$. It obeys the real problem

$$\mu\frac{\partial\Phi^m}{\partial\tau}+\Phi^m+(-1)^m p_1\sqrt{1-\mu^2}\frac{1+\delta_{0,m}}{2}(\Phi^{m-1}+\Phi^{m+1})=\frac{\omega}{2}\int_{-1}^{+1}\Phi^m\chi^m d\mu' \tag{6.90}$$

and boundary conditions (6.89a). The total number of harmonics $M+1$ must be even, so that the number of even and odd Φ^m is equal.

Next, let us expand harmonics χ^m and Φ^m into Legendre polynomial series

$$\begin{aligned}\chi^m(\tau,\mu,\mu') &= \sum_{k=m}^{L_m}\chi_k(\tau)Y^m_k(\mu)Y^m_k(\mu'),\\ \Phi^m(\tau,p,\mu) &= \sum_{k=m}^{L_m}\frac{2k+1}{2}\Phi^m_k(\tau,p)Y^m_k(\mu).\end{aligned} \tag{6.91}$$

Next, let us define a vector of moments $\Phi^m(\tau,p)=[\Phi^m_m,\Phi^m_{m+1},\ldots,\Phi^m_{L_m}]^T$ for the mth harmonic, and a total vector of solution $\Phi=\{\Phi^0,\Phi^1,\ldots,\Phi^M\}^T$. Now we can obtain the system of linear differential equations for the moments $\Phi^m_k(\tau,p)$. To this end, expansions (6.91) are substituted into (6.90), which is then multiplied by $Y^m_p(\mu)$ for $p=m,\ m+1,\ldots,L_m$, $m=0,\ldots,M$, and integrated over μ from -1 to 1. The resulting system of ordinary differential equations for mth harmonic can be written as

$$A^m\frac{d\Phi^m(\tau,p)}{d\tau}+C^m(\tau)\Phi^m(\tau,p)+(-1)^m p_1\frac{1+\delta_{0,m}}{2}((D^{m-1})^T\Phi^{m-1}+D^m\Phi^{m+1})=0 \tag{6.92a}$$

Combining all equations at different m, we can rewrite the system of equations in the following matrix form,

$$\frac{\partial \Phi(\tau,p)}{\partial \tau} + A^{-1}[C(\tau)+p_1 D]\Phi(\tau,p) = 0, \tag{6.92b}$$

where A and C are block-diagonal matrices $A = \mathrm{diag}\,\{A^0, A^1, \ldots, A^M\}$, $C = \mathrm{diag}\,\{C^0, C^1, \ldots, C^M\}$, consisting of standard submatrices A^m and C^m (see section 6.1). Matrix D has the following form:

$$D = \frac{1}{2}\begin{pmatrix} 0 & 2D^0 & & & & & 0 \\ -D^{0T} & 0 & -D^1 & & & & \\ & D^{1T} & 0 & D^2 & & & \\ & & -D^{2T} & 0 & \cdots & & \\ & & & \cdots & \cdots & & D^{M-1} \\ 0 & & & & -(D^{M-1})^T & 0 & \end{pmatrix},$$

with

$$D^m = \begin{pmatrix} d_0 & & & & & 0 & \\ 0 & d_1 & & & & & \\ e_0 & 0 & d_2 & & & & \\ 0 & e_1 & 0 & d_3 & & & \\ \vdots & & \cdots & \cdots & & \cdots & \\ 0 & \cdots & 0 & e_{L_m-m-2} & 0 & & d_{L_m-m} \end{pmatrix},$$

and $d_k = \sqrt{(2m+k+1)(2m+k+2)}$, $e_k = -\sqrt{(k+1)(k+2)}$. For each m in (6.92a), we assume that the last moment is zero, $\Phi^m_{L_m+1} = 0$.

Equation (6.92a) is complemented with Marshak's boundary conditions:

$$\Phi^m_{od}(0) - G^m \Phi^m_{ev}(0) = 0, \tag{6.92c}$$

$$\Phi^m_{od}(\tau_0) + G^m \Phi^m_{ev}(\tau_0) = 2\vec{h}\delta_{0,m}, \quad \vec{h} = (G^0_{1,1}, G^0_{2,1}, G^0_{(L_0+1)/2,1})^T. \tag{6.92d}$$

One of the methods of solving (6.92b) is to assume that the atmosphere consists of n homogeneous layers, each of them having a constant matrix C_i. Then this equation can formally be integrated within the layer to yield:

$$-\Phi_{i-1} + e^{R(\tau_i - \tau_{i-1})}\Phi_i = 0, \;\; R = A^{-1}[C_i + p_1 D], \tag{6.93}$$

where R is a matrix of the layer. Finding matrix exponent $e^{R(\tau_i-\tau_{i-1})}$ is a major part of the solution. Similarly to the 1-D MSH solution (section 6.1), this is done using a singular-value decomposition of matrix R. We also use matrix symmetry to reduce its size by a factor of 2, thus achieving a gain in speed by a factor of approximately 8.

6.6 Similarity transformations

Let us first multiply (6.93) by a block-diagonal matrix $T_0 = \mathrm{diag}\,\{P_0, P_0, \ldots, P_0\}$, where P_0 is a unitary matrix of dimension L_0 sorting a vector into even and odd components. Matrix T_0 will transform vector Φ into a vector $y_1 = [(\Phi^0_t, \Phi^0_b), \ldots, (\Phi^M_t, \Phi^M_b)]^T$, in which Φ^m_t and Φ^m_b consist of even and odd elements of vector Φ^m if m is even, and vice versa if m is odd. Matrix R will turn into a matrix G:

$$G = T_0 R T_0^T = T_0 A^{-1} T_0^T [T_0 C_i T_0^T + p_1 T_0 D T_0^T],$$

where each submatrix of A^{-1}, C and D, corresponding to the mth harmonic, will be changed according to the rule $X^m \to P_0 X^m P_0^T$. Specifically, diagonal matrix C^m will be sorted into even and odd components,

$$P_0 C^m P_0^T = \begin{pmatrix} C_t^m & 0 \\ 0 & C_b^m \end{pmatrix}.$$

Matrix D^m will be sorted accordingly,

$$P_0 D^m P_0^T = \begin{pmatrix} D_t^m & 0 \\ 0 & D_b^m \end{pmatrix},$$

where

$$D_t^m = \begin{pmatrix} d_m^m & & & 0 \\ e_m^m & d_{m+2}^m & & \\ 0 & e_{m+2}^m & \cdots & \\ 0 & 0 & e_{L_m-3}^m & d_{L_m-1}^m \end{pmatrix},$$

and

$$D_b^m = \begin{pmatrix} d_{m+1}^m & & & 0 \\ e_{m+1}^m & d_{m+3}^m & & \\ 0 & e_{m+3}^m & \cdots & \\ 0 & 0 & e_{L_m-2}^m & d_{L_m}^m \end{pmatrix}.$$

Matrix $(A^m)^{-1}$ will be transformed into a matrix

$$P_0 (A^m)^{-1} P_0^T = \begin{pmatrix} 0 & A_1^m \\ A_1^{mT} & 0 \end{pmatrix}.$$

Matrices A_1^m can be found analytically without matrix inversion due to a simple structure of A^m. For example, for $L_0 = 6$,

$$A_1^m = \begin{pmatrix} a_1^{-1} & \frac{-a_2}{a_1 a_3} & \frac{a_2 a_4}{a_1 a_3 a_5} \\ 0 & a_3^{-1} & \frac{-a_4}{a_3 a_5} \\ 0 & 0 & a_5^{-1} \end{pmatrix},$$

where elements a_i of matrix A^m were defined in section 6.1.

Next, we apply a unitary transformation $\hat{P}$ to the modified equation (6.93). Matrix $\hat{P}$ is designed to sort vector $y_1 = [(\Phi_t^0, \Phi_b^0), \ldots, (\Phi_t^M, \Phi_b^M)]^T$ into the top and bottom parts

$$y = \hat{P} y_1 = [(\Phi_t^0, \Phi_t^1, \ldots, \Phi_t^M),\ (\Phi_b^0, \Phi_b^1, \ldots, \Phi_b^M)]^T = (y_t, y_b)^T. \tag{6.94}$$

Therefore, $\hat{P}$ is just a sorting matrix of size $(M+1) \times (M+1)$, having as its elements the following submatrices of dimension $(L_0/2) \times (L_0/2)$: identity matrices I in place of units, and zero matrices in place of zeroes. For example, for $M = 3$,

$$\hat{P} = \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & 0 & I \end{pmatrix}.$$

These similarity transformations bring (6.93) to the form:

$$-\begin{pmatrix} y_t \\ y_b \end{pmatrix}_{i-1} + \exp(F\Delta\tau_i) \times \begin{pmatrix} y_t \\ y_b \end{pmatrix}_i = 0, \tag{6.95}$$

where y_t, y_b are defined by (6.94),

$$F = \begin{pmatrix} 0 & B \\ T & 0 \end{pmatrix},$$

and matrices T, B have the form

$$T = \begin{pmatrix} A_1^{0T}C_t^0 & p_1 A_1^{0T} D_t^0 & 0 & \cdots & 0 \\ -p_1 A_1^{1T} D_t^{0T}/2 & A_1^{1T} C_t^1 & -p_1 A_1^{1T} D_t^1/2 & & \vdots \\ 0 & p_1 A_1^{2T} D_t^{1T}/2 & A_1^{2T} C_t^2 & \cdots & 0 \\ \vdots & & \cdots & \cdots & p_1 (A_1^{M-1})^T D_t^{M-1}/2 \\ 0 & \cdots & 0 & -p_1 A_1^{MT} (D_t^{M-1})^T/2 & A_1^{MT} C_t^M \end{pmatrix},$$

$$B = \begin{pmatrix} A_1^0 C_b^0 & p_1 A_1^0 D_b^0 & 0 & \cdots & 0 \\ -p_1 A_1^1 D_b^{0T}/2 & A_1^1 C_b^1 & -p_1 A_1^1 D_b^1/2 & & \vdots \\ 0 & p_1 A_1^2 D_b^{1T}/2 & A_1^2 C_{bt}^2 & \cdots & 0 \\ \vdots & & \cdots & \cdots & p_1 A_1^{M-1} D_b^{M-1}/2 \\ 0 & \cdots & 0 & -p_1 A_1^M (D_b^{M-1})^T/2 & A_1^M C_b^M \end{pmatrix}.$$

6.6.1 Singular value decomposition

Matrix F possesses a symmetry necessary to decrease its size by a factor of 2 in the SVD problem. Suppose that SVD was found for matrix F, which is our ultimate purpose:

$$F = Z\Lambda_1 Z^{-1}, \tag{6.96}$$

where Z is a matrix of eigenvectors, and Λ_1 is a diagonal matrix of eigenvalues. Consider matrix

$$F^2 = \begin{pmatrix} BT & 0 \\ 0 & TB \end{pmatrix} = Z\Lambda_1^2 Z^{-1}, \tag{6.97}$$

which is block-diagonal, has the same eigenvectors as matrix F and has squared eigenvalues of matrix F. Matrices BT and TB have the same eigenvalues, and their matrices of eigenvectors are related by a linear matrix transformation. So, it is sufficient to find the SVD for only one of them, for example $BT = U\Lambda^2 U^{-1}$. Then, eigenvalues of F are $\pm\sqrt{\lambda_i^2}$, where λ_i^2 are eigenvalues of both BT and of F^2, and

$$\Lambda_1 = \begin{pmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{pmatrix}.$$

After that, we can reconstruct matrix Z as

$$Z = \begin{pmatrix} U & U \\ V & V_- \end{pmatrix},$$

where unknown matrices V and V_- correspond to positive and negative eigenvalues. These matrices can be found from SVD (6.96) $FZ = Z\Lambda_1$, which gives: V_-=-V, and $V = TU\Lambda^{-1}$. In order to find the SVD for matrix BT, we use QR algorithm with a preliminary reduction of matrix to the Hessenberg form.

6.6.2 Solution for moments

Now we can find the matrix exponential of (6.95) explicitly:

$$\exp\left\{\begin{pmatrix} 0 & B \\ T & 0 \end{pmatrix}\Delta\tau_i\right\} = Z\exp\left\{\begin{pmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{pmatrix}\Delta\tau_i\right\}Z^{-1}, \text{ where}$$
$$Z = \begin{pmatrix} U & U \\ V & -V \end{pmatrix},\ Z^{-1} = \frac{1}{2}\begin{pmatrix} U^{-1} & V^{-1} \\ U^{-1} & -V^{-1} \end{pmatrix}. \tag{6.98}$$

With this decomposition and multiplication of (6.95) by $2SZ^{-1}$, where the scaling transformation S is defined as

$$S = \begin{pmatrix} 0 & I \\ e^{-\Lambda\Delta\tau} & 0 \end{pmatrix},$$

we obtain

$$-S_i Z_i^{-1} y_{i-1} + H_i Z_i^{-1} y_i = 0,\ i = 1, 2, \ldots, n, \tag{6.99}$$

where

$$H = S\exp\left\{\begin{pmatrix} \Lambda & 0 \\ 0 & -\Lambda \end{pmatrix}\Delta\tau_i\right\} = \begin{pmatrix} 0 & e^{-\Lambda\Delta\tau} \\ I & 0 \end{pmatrix}.$$

At high M and L_0, when the real parts of maximal eigenvalues are large, the scaling transformation S stabilizes system of equations (6.99) by extinguishing positive exponents in its coefficients.

Finally, let us denote $E_i = \exp(-\Lambda\Delta\tau_i)$ and rewrite (6.99) as follows:

$$-U^{-1}y_{i-1}^t + V^{-1}y_{i-1}^b + E_i U^{-1}y_i^t - E_i V^{-1}y_i^b = 0 \tag{6.100a}$$

$$-E_i U^{-1}y_{i-1}^t - E_i V^{-1}y_{i-1}^b + U^{-1}y_i^t + V^{-1}y_i^b = 0. \tag{6.100b}$$

This set of equations along with boundary conditions (6.92) provides a closed system of linear equations in n atmospheric layers, which is efficiently solved with the Gauss elimination method.

6.6.3 Solution for the OTF

Once moments are found, the OTF can be reconstructed by summing series (6.91) and (6.88). Similarly to the 1-D case, truncation of these series at M and L_0 leads to the errors in solution, which appear as non-physical oscillations of the OTF in zenith and azimuthal angles as well as in parameter p. To compensate for these errors, we use a smoothing method of integration of the source function. In order to describe the algorithm, let us integrate (6.87) from τ_i to τ_{i+1} to obtain a relation between OTF values at the interfaces of ith layer:

$$\Psi_i(\tau, p; \mu, \varphi) = e^{-\frac{\alpha\Delta\tau_i}{\mu}}\Psi_{i+1}(\tau, p; \mu, \varphi) + \frac{1}{\mu}\int_{\tau_i}^{\tau_{i+1}} e^{-\frac{\alpha(\tau-\tau_i)}{\mu}} J(\tau, p; \mu, \varphi)\, d\tau, \tag{6.101}$$

where $\alpha = 1 - ip_1\sqrt{1-\mu^2}\cos\varphi$, and $J(\tau, p; \mu, \varphi)$ is a right-hand side of (6.87),

$$J(\tau,p;\mu,\varphi)=\omega(\tau)\sum_{m=0}^{M}\frac{2-\delta_{0,m}}{2}\cos m\varphi\sum_{k=m}^{L_m}\chi_k(\tau)\psi_k^m(\tau,p)Y_k^m(\mu). \qquad (6.101a)$$

The integral in the right-hand side of (6.101) contains integrals from the moments

$$\int_{\tau_i}^{\tau_{i+1}} e^{-\frac{\alpha(\tau-\tau_i)}{\mu}}\Phi_k^m\,d\tau.$$

To find these terms, multiply equation for moments (6.92a) for each m by $e^{-[\alpha(\tau-\tau_i)/\mu]}$ and integrate it within a homogeneous layer from τ_i to τ_{i+1}:

$$A^m\left\{e^{-\frac{\alpha(\tau_{i+1}-\tau_i)}{\mu}}(\Phi^m)_{i+1}-(\Phi^m)_i\right\}+\left\{\frac{\alpha}{\mu}A^m+C_i^m\right\}\int_{\tau_i}^{\tau_{i+1}}e^{-\frac{\alpha(\tau-\tau_i)}{\mu}}\Phi^m\,d\tau$$
$$+(-1)^m p_1\frac{1+\delta_m}{2}\left\{(D^{m-1})^T\int_{\tau_i}^{\tau_{i+1}}e^{-\frac{\alpha(\tau-\tau_i)}{\mu}}\Phi^{m-1}\,d\tau+D^m\int_{\tau_i}^{\tau_{i+1}}e^{-\frac{\alpha(\tau-\tau_i)}{\mu}}\Phi^{m+1}\,d\tau\right\}$$
$$=\ 0; \qquad (6.102)$$

Next, multiply (6.102) by $(A^m)^{-1}$ and denote

$$S^m=\int_{\tau_i}^{\tau_{i+1}}e^{-\frac{\alpha(\tau-\tau_i)}{\mu}}\Phi^m\,d\tau,\ \beta^m=e^{-\frac{\alpha(\tau_{i+1}-\tau_i)}{\mu}}(\Phi^m)_{i+1}-(\Phi^m)_i.$$

Defining vectors $S=\{S^0,S^1,\ldots,S^M\}$, and $\beta=\{\beta^0,\beta^1,\ldots,\beta^M\}$, we can rewrite the above equation as

$$(\frac{\alpha}{\mu}+R)S=-\beta, \qquad (6.103)$$

where R is a matrix of layer introduced by formula (6.93). Thus, we have obtained a linear algebraic system of equations for the integral of the moments S, which can be solved, for example, using the Gauss elimination method. In practice, we first apply similarity transformations in order to use the SVD found before. This allows us to substitute operation of matrix inversions for each atmospheric layer and each viewing direction by an operation of matrix multiplication. As a result, calculation of the OTF in arbitrary directions takes only a small part of the total computer time even at a considerable number of view angles.

The smoothing procedure described ensures a rapid convergence of the solution with the increase of orders of azimuthal and Legendre polynomial expansions. As a result, acceptable accuracy of about 1% at $|\mu|\geq 0.3$, $p\leq 3$–10 is already achievable in a wide range of atmospheric conditions at $M=9$–15, $L_0=23$–35. The specified range of zenith view angles satisfies virtually all possible applications of the optical transfer function, and a wider range of spatial frequencies may be required only in the presence of extremely dense and geometrically thin aerosol layers.

Fig. 6.3 illustrates the behavior of the amplitude and phase of the OTF, and of the spherical albedo of the atmosphere in different atmospheric and viewing conditions.

The algorithm developed works reliably for low and medium spatial frequencies, $p\leq 2$–10, depending on aerosol stratification in the atmosphere. At higher p,

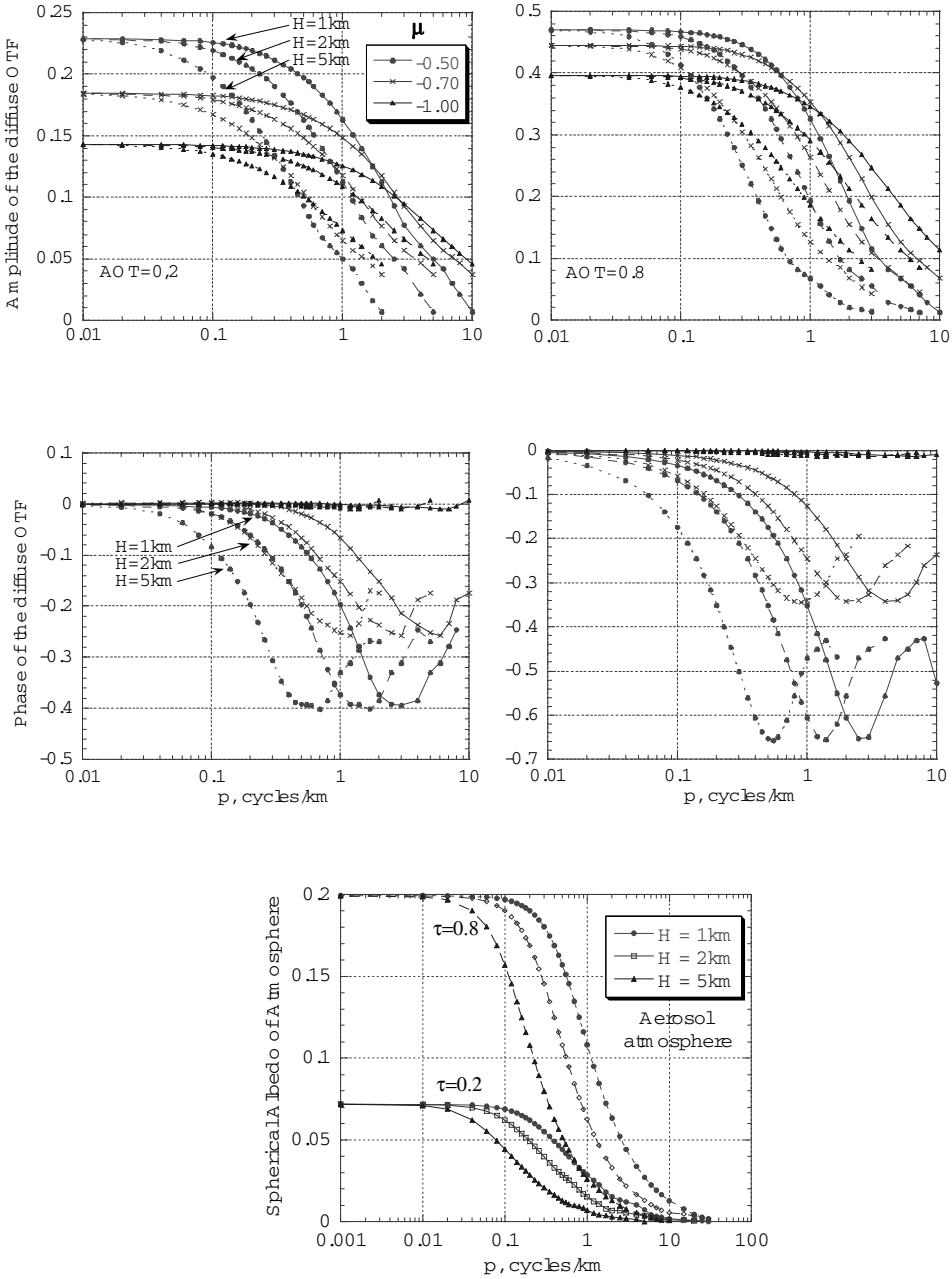


Fig. 6.3. Amplitude and phase of OTF and spherical albedo of atmosphere as a function of spatial frequency (from Lyapustin and Muldashev, 2001). The results shown were obtained for the continental aerosol atmosphere for clear ($\tau = 0.2$) and hazy ($\tau = 0.8$) conditions. The stark difference of 3-D solution from 1-D values is the dependence of the 3-D result on the vertical profile. Here, calculations are shown for three heights of aerosol layer – 1, 2, and 5 km.

oscillations of the complex term degrade convergence of the algorithm. On the other hand, the multiple scattering fraction of the solution quickly decreases with the increase of p. For this reason, at $p > p_{ss} = 2$ the OTF is computed in the single scattering as was proposed by Ioltukhovskii (1999). Analysis of the full solution (6.87) for a variety of scenes shows that the error due to this approximation is negligible.

6.7 Code SHARM-3D

The described development of the 3-D Green's function method and of the numerical solution for the OTF led to the creation of code SHARM-3D (Lyapustin and Wang, 2005). This code performs fast and accurate simulations of the monochromatic radiance at the top of atmosphere over spatially variable surfaces with Lambertian or anisotropic reflectance. The atmosphere is assumed to be laterally uniform across the image, and consisting of two layers with aerosols contained in the bottom layer. Such a model plays an important role in the space-borne remote sensing of atmospheric aerosol and surface reflectance. Far from localized sources, aerosols vary on a much coarser scale than the surface reflectance, typically at 50–60 km (Anderson et al., 2003). Within such distances, the aerosol amount in the atmospheric column and its radiative properties can often be considered approximately uniform, and all of the spatial and angular variability of the measured signal can be attributed to the variable surface reflectance. The GF method developed offers a particular advantage when the atmosphere is laterally homogeneous, because the atmospheric radiative functions need to be obtained only once for the whole area of interest.

6.7.1 Parameterized SHARM-3D solution

The Green's function solution for the top-of-atmosphere radiance was given earlier by equations (6.85)–(6.86). Formula (6.84) is used to compute the albedo of a spatially heterogeneous surface, and equation (6.42) is used to compute the average surface-reflected radiance. Several parameterizations discussed earlier greatly simplify the radiative transfer algorithm and gain a speed advantage of up to a factor of 10^3 as compared to SHDOM, yet retaining the high accuracy, which is generally better than 1% (Lyapustin, 2002).

Further improvement in speed is achieved by the use of the LSRT model which parameterizes the GF-solution into a weakly nonlinear function of spatially variable kernel weights with angular-dependent coefficients that need to be calculated only once for the whole image. Using functions (6.49)–(6.54) introduced earlier, the second and third terms of (6.85) (the surface-reflected radiance directly transmitted at TOA, and the mean surface-reflected radiance diffusely transmitted at TOA) can be written as follows:

$$\begin{aligned} L_s(r;s_0,s) \;&\cong\; S_\lambda \mu_0\, e^{-\tau(\frac{1}{\mu_0}+\frac{1}{|\mu|})}\{\rho(r;\mu_0,\mu,\varphi) + \alpha c_0 \rho_1(r;\mu)\bar{\rho}_2(\mu_0)\} \\ &+\; e^{-\tau/|\mu|}\alpha[k^L(r)E_0^d(\mu_0) + k^G(r)D_G^1(\mu_0,\mu,\varphi) + k^V(r)D_V^1(\mu_0,\mu,\varphi)]; \end{aligned} \tag{6.104}$$

$$\begin{aligned}\bar{L}_s^d(\mu_0,\mu,\varphi) &= S_\lambda \mu_0\, e^{-\tau/\mu_0}\{[\bar{k}^L G^{av}(\mu)+\bar{k}^G G_G^1(\mu_0,\mu,\varphi)+\bar{k}^V G_V^1(\mu_0,\mu,\varphi)]\\ &+ \alpha c_0[\bar{k}^L G^{av}(\mu)+\bar{k}^G G_G^{11}(\mu)+\bar{k}^V G_V^{11}(\mu)]\bar{\rho}_2(\mu_0)\}\\ &+ \alpha[\bar{k}^L E_0^d(\mu_0)G^{av}(\mu)+\bar{k}^G H_G^1(\mu_0,\mu,\varphi)+\bar{k}^V H_V^1(\mu_0,\mu,\varphi)].\end{aligned} \tag{6.105}$$

The surface albedo is written as:

$$\begin{aligned}q(r;\mu_0) &= E_0^{-1}(\mu_0)\{\mu_0\, e^{-\tau/\mu_0} q_2(r;\mu_0)+k^L(r)E_0^d(\mu_0)\\ &+ k^G(r)D_G^3(\mu_0)+k^V(r)D_V^3(\mu_0)\}.\end{aligned} \tag{6.106}$$

The algorithm described uses two different solutions of 1-D code SHARM. The first one is a standard solution with the atmosphere illuminated at the top. It provides the path radiance at TOA along with its azimuthal harmonics at the bottom of the atmosphere, surface irradiance, atmospheric transmittance and spherical albedo. In the second solution, the atmosphere is illuminated from the bottom; in other words, it corresponds to a reversed order of atmospheric layers. This solution provides azimuthal harmonics of the Green's function in the multiple scattering. Because aerosol scattering may cause the Green's function to be very asymmetric in the aureole region, the harmonics of the single-scattering term are calculated separately using the high-order Gaussian quadrature for azimuthal angle ($N_{SS} = 129$). With this separation, a relatively low order of MSH ($nb = 24$–36) can be used in the multiple scattering calculations. This approach reduces the overall computing time and preserves the total accuracy. The zenith angle integration uses Gaussian quadrature of the order $N_q = nb/2 + 10$. Because the Legendre polynomial of the order $2N$ is integrated exactly with the quadrature of the order N, and the kernels f_G, f_V can be approximated by the low-order polynomials, the quadrature N_q warrants accurate integration.

Computing the last term of (6.85) requires knowledge of the atmospheric OTF. Because computing OTF for given atmospheric conditions consumes a lion's share of the time required by the SHARM-3D algorithm, we pre-computed a look-up table (LUT) of certain functions and developed an algorithm which reconstructs the full OTF from the LUT based on the symmetry and scaling properties of the OTF. The detail of the OTF LUT algorithm can be found in (Lyapustin and Wang, 2005).

With the OTF known, computing the last variational term of (6.85) requires two 2-D Fourier transforms: the first (direct) transform calculates spatial spectrum of variation of surface albedo, and the second (inverse) transform restores the spatial variation of the TOA radiance. These operations are efficiently performed with the FFT algorithm (Press et al., 1992) that takes $N \log_2 N$ operations for the image with N^2 pixels. The selected Fourier-transform approach based on OTF is more efficient than the approach with spatial integration of the atmospheric PSF that requires $\sim N^2$ operations. In summary, for small images the computing time is entirely defined by the 1-D radiative transfer calculations of code SHARM. For large images, the overall time is affected by calculation of the last variational term of (6.85).

The use of the linear LSRT model and of the pre-computed LUT of the OTF raises the computational speed of simulating the Earth's outgoing spectral radiance

to its practical limit, when any further speed improvements can only be of an incremental nature.

6.8 Discussion

Similarly to the 1-D code SHARM, the input files for SHARM-3D include a configuration file and files of atmospheric and surface properties. The details on the input parameters are provided in the document 'SHARM Manual'.

The LSRT BRF model is used to simulate anisotropic reflectance of land surfaces. In order to realistically model reflectance of inland rivers and lakes, especially in the glint area, we are using the model of Nakajima and Tanaka (1983). The wind speed is assumed constant across the image. To approximately model reflectance of suspended matter or reflectance from the shallow bottom, a variable Lambertian offset can be specified for the water pixels.

Because it neglects the nonlinear interactions in the variation of surface reflectance, code SHARM-3D should not be used at resolutions or scales of inhomogeneity typical of establishing the 1-D regime, e.g. 3–5 km. In these conditions, SHARM-3D produces systematic biases for pixels that are brighter or darker than the average. The magnitude of biases, which tend to reduce the contrast of calculated radiance, may reach several percent. This is easy to demonstrate using an example of Lambertian surface. Over the large homogeneous areas of the image, the multiple reflections from the surface enhance the surface-reflected radiance proportionally to the albedo of this particular area, $(1-q^{area}c_0)^{-1}$. At the same time, the linearized solution (6.85) calculates this enhancement as proportional to the albedo averaged over the whole image, $(1-\bar{q}c_0)^{-1}$, which leads to the biases mentioned.

To allow the user make calculations in the 1-D regime, we implemented an independent pixel approximation (IPA) as a separate mode of calculations. The mode is specified in the surface properties file (*.sfc) by parameter *szDimRT*, which can take a value of either '1D' or '3D'. The algorithm behind the IPA is a rigorous 1-D Green's function method (equations (6.33)–(6.36)) which has the same accuracy as code SHARM, but becomes progressively faster with the increase of the image size. The IPA algorithm does not use parameterizations. Instead, it achieves a rigorous convergence for the series of multiple reflections for each surface pixel. For this reason, the '1D' mode is considerably slower than the '3D' mode.

A relative accuracy of SHARM-3D and IPA with respect to SHDOM is shown in Fig. 6.4 for a particular surface realization from the Landsat-7 scene for the Oklahoma site. At the time of acquisition (April 4, 2000), the scene's primary composite elements were the plowed fields (dark) and spring grasses (bright). In this test, we used the image subset of 64×64 pixels in the near-IR band 4 in order to generate realistic surface reflectance with BRF shape assigned according to the surface brightness (Lyapustin, 2002). The simulations were performed at 30 m resolution in clear atmospheric conditions ($\tau = 0.2$) in the near-IR channel. The nadir TOA radiance calculated by SHARM-3D is shown on the left. The middle image shows the relative difference between SHARM-3D and SHDOM calculations. One can see that the agreement of these codes is better than $\sim$1% for over 90% of the pixels. For comparison, the right image shows that the equivalent 1-D solution (SHDOM – IPA)/ SHDOM has the range of errors of $\pm$17%.

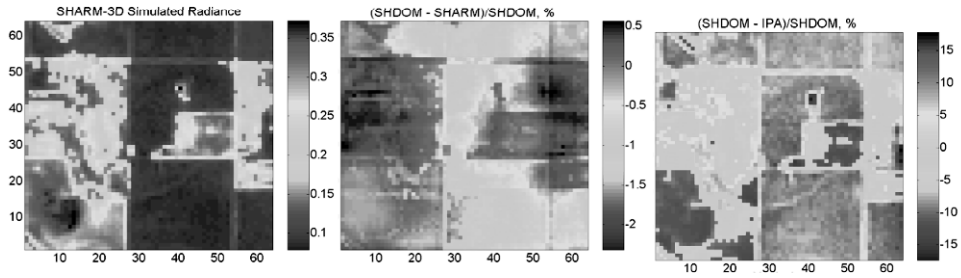


Fig. 6.4. Comparison of code SHARM-3D with SHDOM and independent pixel approximation (IPA) (from Lyapustin, 2002). The left image shows SHARM-3D simulated TOA radiance. The 64×64 pixels surface area at 30-m resolution was modeled using Landsat-7 reflectance in band 4 over the Oklahoma site on April 4, 2000. The middle and right images show the relative difference of SHARM-3D and IPA solutions with SHDOM. [For the color version of this figure, see the color plate section.]

As a brief summary, let us list the main features of code SHARM-3D. It was developed for rapid simulations of TOA radiance over spatially variable surfaces with Lambertian or anisotropic reflectance with high accuracy of $\sim$1%. The code also computes surface albedo corresponding to given solar zenith angle, atmospheric conditions, and spatial distribution of BRF. The surface boundary condition is periodic. The atmosphere is laterally uniform, and consists of two vertical slabs with aerosols in the bottom layer. Code SHARM-3D performs simultaneous calculations for all specified incidence-view geometries, and multiple wavelengths in one run. The range of view zenith angles is at present limited by the maximal value of $\mu = -0.3$ ($\theta \leq 72.5^\circ$) in the look-up table of the pre-computed OTF. Also, the maximal LUT value of the total optical thickness of atmosphere is 0.9. If $\tau > 0.9$, we assume that $\mathrm{OTF}(\tau) = \mathrm{OTF}(0.9)$. This assumption has little impact on the accuracy of radiance calculations because (1) all other terms are calculated accurately, (2) the path radiance dominates the total TOA radiance at high optical thickness, and (3) the relative contribution of the variational term decreases with τ proportionally to the decrease of the surface irradiance.

Acknowledgments

Drs Lyapustin and Wang are grateful to NASA EOS program and in particular, NASA Terrestrial Ecology Program (Dr D. Wickland) for support of this research. Dr Lyapustin would like to express his special gratitude to Dr Robert E. Murphy for his long-term help and support.

References

Anderson, T. L., R. J. Charlson, D. M. Winker, J. A. Ogren, Kim Holmenm, 2003: Mesoscale variations of tropospheric aerosols, *J. Atm. Sci.*, **60**, 119–136.

Bell, G. I., and S. Glasstone, 1970: *Nuclear Reactor Theory.* Van Nostrand Reinhold, New York.

Benassi, M., R. D. M. Garcia, A. Karp et al., 1984: A high-order spherical harmonics solution to the standard problem in radiative transfer, *Astrophys. J.*, **280**, 2, 853–864.

Bodhaine, B. A., N. B. Wood, E. G. Dutton, J. R. Slusser, 1999: On Rayleigh Optical Depth Calculations, *J. Atmos. Oceanic Technology*, **16**, 1854–1861.

Cahalan R. F., et al., 2005: The international intercomparison of 3D radiation codes (I3RC). Bringing together the most advanced radiative transfer tools for cloudy atmospheres, *Bull. Amer. Meteor. Soc.*, **86**, 1275–1293.

Case, K. M., and P. F. Zweifel, 1967: *Linear Transport Theory.* Addison-Wesley Publishing Company, Reading, MS.

Chandrasekhar, S., 1958: On the diffuse reflection of a pencil of radiation by a plane-parallel atmosphere, *Proc. Nat. Acad. Sci.*, **44**, 933–940.

Chandrasekhar, S., 1960: *Radiative Transfer.* Dover, New York.

Cox, C., and W. Munk, 1954: Measurements of the roughness of the sea surface from photographs of the Sun's glitter, *J. Opt. Society Am.,* **44**, 838–850.

Dave, J. V., 1975: A direct solution of the spherical harmonics approximation to the radiative transfer equation for an arbitrary solar elevation, *J. Atmos. Sci.*, **32**, 790–798.

Dave, J. V., and B. H. Armstrong, 1974: Smoothing of intensity curve obtained from a solution of spherical harmonics approximation to transfer equation, *J. Atmos. Sci.*, **31**, 1934–1937.

Davis, A. B., I. N. Polonski, and A. Marshak, 2009: Space-time Green functions for diffusive radiation transport, in application to active and passive cloud probing. In *Light Scattering Reviews, 4. Single Light Scattering and Radiative Transfer*, A. A. Kokhanovsky (ed.), Springer, 169–292.

Diner, D. J., and J. V. Martonchik, 1984a: Atmospheric transfer of radiation above an inhomogeneous non-Lambertian reflective ground – I. Theory, *J. Quant. Spectrosc. Radiat. Transfer*, **31**, 97–125.

Diner, D. J., and J. V. Martonchik, 1984b: Atmospheric transfer of radiation above an inhomogeneous non-Lambertian reflective ground – II. Computational considerations and results, *J. Quant. Spectrosc. Radiat. Transfer*, **32**, 279–304.

Diner, D. J., and J. V. Martonchik, 1985: Influence of aerosol scattering on atmospheric blurring of surface features, *IEEE Trans. Geosci. Remote Sens.*, **GE-23**, 618–624.

Dubovik, O., A. Sinyuk, T. Lapyonok, B. N. Holben, M. Mishchenko, P. Yang, T. F. Eck, H. Volten, O. Munoz, B. Veihelmann, van der Zander, J.-F. Leon, M Sorokin, and I. Slutsker, 2006: Application of spheroid models to account for aerosol particle non-sphericity in remote sensing of desert dust, *J. Geophys. Res.*, **111**, D11208, doi:10.1029/2005JD006619.

Evans, K. F., 1998: The spherical harmonics discrete ordinate method for three-dimensional atmospheric radiative transfer, *J. Atm. Sci.,* **55**, 429–446.

Gatebe, C., M. D. King, A. I. Lyapustin, G. T. Arnold, and J. Redermann, 2005: Airborne spectral measurements of ocean directional reflectance, *J. Atmos. Sci.*, **62**, 1071–1091.

Germogenova, T. A., 1986: *The Local Properties of the Solution of the Transport Equation* (in Russian). Nauka, Moscow, Russia.

Gerstl, S. A. W., 1982: Application of the adjoint method in atmospheric radiative transfer calculations. in *Atmospheric Aerosols*: *Their Formation, Optical Properties and Effects.* A. Deepak (ed.). Spectrum Press, Hampton VA, 241–254.

Hansen, J. E., and J. W. Hovenier, 1971: The doubling method applied to multiple scattering of polarized light, *J. Quant. Spectrosc. Radiat. Transfer*, **11**, 809–812.

Holben, B. N., T. F. Eck, I. Slutsker, D. Tanré, J. P. Buis, A. Setzer, E. Vermote, J. A. Reagan, Y. J. Kaufman, T. Nakajima, F. Lavenu, I. Jankowiak, A. Smirnov, 1998: AERONET-A Federated instrument network and data archive for aerosol characterization, *Rem. Sens. Environ.*, **66,** 1–16.

Ioltukhovskii, A. A., 1999: Radiative transfer over the surface with an arbitrary reflection: Green's functions method, *Transport Theory and Statistical Physics*, **28** (4), 349–368.

Karp, A. H., 1981: Computing the angular dependence of the radiation of a planetary atmosphere, *J. Quant. Spectrosc. Radiat. Transfer*, **25**, 403–412.

Karp, A. H., J. Greenstadt and J. A. Fillmore, 1980: Radiative transfer through an arbitrarily thick scattering atmosphere, *J. Quant. Spectrosc. Radiat. Transfer*, **24**, 391–406.

King, M., and R. Greenstone (eds), 1999: *EOS Reference Handbook: A Guide to Earth Science Enterprise and the Earth Observation System* (p. 355). Greenbelt, MD: EOS Project Science Office, NASA/Goddard Space Flight Center.

Kneizys, F. X., L. W. Abreu, G. P. Anderson, J. H. Chetwynd, E. P. Shettle, A. Berk, L. S. Bernstein, D. C. Robertson, P. Acharya, L. S. Rothman, J. E. A. Selby, W. O. Gallery, S. A. Clough, 1996: *The MODTRAN 2/3 Report and LOWTRAN 7 Model, Modtran Report*. Ontar Corporation, North Andover, MA.

Landgraf J., Hasekamp O. P., Trautmann T., 2002: Linearization of radiative transfer with respect to surface properties, *J. Quant. Spectrosc. Radiat. Transfer*, **72** (4), 327–339.

Lenoble, J. (ed.), 1985: *Radiative Transfer in Scattering and Absorbing Atmospheres: Standard Computational Procedures*. A. Deepak Publishing, Hampton, VA.

Lenoble, J., M. Herman, J.L. Deuzé, B. Lafrance, R. Santer, D. Tanre, 2007: A successive order of scattering code for solving the vector equation of transfer in the earth's atmosphere with aerosols, *J. Quant. Spectrosc. Radiat. Transfer*, **107**, 479–507.

Lucht W., Schaaf C. B., Strahler A. H., 2000: An algorithm for the retrieval of albedo from space using semiempirical BRDF models, *IEEE Trans. Geosci. Remote Sens.,* **38**, 977–998.

Lyapustin, A. I., 2001: 3-D effects in the remote sensing of surface albedo, *IEEE Trans. Geosci. Remote Sensing*, **39**, 254–263.

Lyapustin, A., 2002: Radiative transfer code SHARM-3D for radiance simulations over a non-Lambertian nonhomogeneous surface: intercomparison study, *Appl. Optics*, **41,** 5607–5615.

Lyapustin, A., 2005: Radiative transfer code SHARM for atmospheric and terrestrial applications, *Appl. Optics*, **44**, 7764–7772.

Lyapustin, A. I., and Y. J. Kaufman, 2001: Role of adjacency effect in the remote sensing of aerosol. *J. Geophys. Res.*, **106**, D11, 11,909–11,916.

Lyapustin, A., and Yu. Knyazikhin, 2001: Green's function method in the radiative transfer problem. I: Homogeneous non-Lambertian surface, *Appl. Optics*, **40**, 3495–3501.

Lyapustin, A., and Yu. Knyazikhin, 2002: Green's function method in the radiative transfer problem. II: Spatially heterogeneous anisotropic surface, *Appl. Optics*, **41**, 5600–5606.

Lyapustin, A. I., and T. Z. Muldashev, 1999: Method of spherical harmonics in the radiative transfer problem with non-Lambertian surface, *J. Quant. Spectrosc. Radiat. Transfer*, **61**, 545–555.

Lyapustin, A. I., and T. Z. Muldashev, 2000: Generalization of Marshak boundary condition for non-Lambert reflection, *J. Quant. Spectrosc. Radiat. Transfer*, **67**, 457–464.

Lyapustin, A. I., and T. Z. Muldashev, 2001: Solution for atmospheric optical transfer function using spherical harmonics method. *J. Quant. Spectrosc. Radiat. Transfer*, **68**, 43–56.

Lyapustin, A. I., and J. L. Privette, 1999: A new algorithm for retrieving surface BRDF from ground measurements: atmospheric sensitivity study, *J. Geophys. Res.*, **104**, 6257–6268.

Lyapustin, A., and Y. Wang, 2005: Parameterized code Sharm-3D for radiative transfer over inhomogeneous surfaces, *Appl. Optics*, **44**, 7602–7610.

Lyapustin, A., Wang, Y., 2009: The time series technique for aerosol retrievals over land from MODIS, pp. 69–99, in *Satellite Aerosol Remote Sensing over Land*, A. Kokhanovky and G. de Leeuw (eds). Springer Praxis Books, IBSN: 978-3-540-69396-3.

Marchuk, G., G. Mikhailov, N. Nazaraliev, R. Darbinjan, B. Kargin, and B. Elepov, 1980: *The Monte Carlo Methods in Atmospheric Optics.* Springer-Verlag, New York.

Marshak R. E., 1947: Note on the spherical harmonics method as applied to the Milne problem for a sphere, *Phys. Rev.*, **71**, 443–446.

Martonchik, J. V., D. J. Diner, B. Pinty, M. M. Verstratete, R. B. Myneni, Yu. Knyazikhin, and H. R. Gordon, 1998: Determination of land and ocean reflective, radiative and biophysical properties using multiangle imaging, *IEEE Trans. Geosci. Remote Sens.,* **36**, 1266–1281.

Mayer, B., and A. Kylling, 2005: Technical note: The libRadtran software package for radiative transfer calculations: Description and examples of use. *ACPD,* **5**, 1319–1381.

Mekler, Yu., and Y. J. Kaufman, 1980: The effect of Earth's atmosphere on contrast reduction for a non-uniform surface albedo and 'two-halves' field, *J. Geophys. Res.*, **85**, 4067–4083.

Muldashev, T. Z., A. I. Lyapustin, and U. M. Sultangazin, 1999: Spherical harmonics method in the problem of radiative transfer in the atmosphere-surface system, *J. Quant. Spectrosc. Radiat. Transfer*, **61**, 393–404.

Nakajima, T., and M. Tanaka, 1983: Effect of wind-generated waves on the transfer of solar radiation in the atmosphere–ocean system, *J. Quant. Spectrosc. Radiat. Transfer,* **29**, 521–537.

Otterman, J., and R. S. Fraser, 1979: Adjacency effects on imaging by surface reflection and atmospheric scattering: Cross radiance to zenith, *Appl. Opt.*, **18**, 2852–2860.

Pearce, W. A., 1977: *A study of the effect of atmosphere on the Thematic Mapper observations.* Rep. 004-77, EG&G/Washington Anal. Serv. Center, Appl. Syst. Dept., Riverdale, MD.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1992: *Numerical Recipes in C*, 2nd edn, Cambridge University Press, New York.

Qin Y., Box M. A., 2005: Analytic green's function for radiative transfer in plane-parallel atmospheres, *J. Quant. Spectrosc. Radiat. Transfer*, **62** (8), 2910–2924.

Rahman, H., B. Pinty, and M. M. Verstraete, 1993: Coupled surface–atmosphere reflectance (CSAR) model. 2. Semiempirical surface model usable with NOAA advanced very high resolution radiometer data, *J. Geophys. Res.,* **98**, 20,791–20,801.

Ricchiazzi, P., S. Yang, C. Gautier, and D. Sowle, 1998: SBDART: A research and teaching software tool for plane-parallel radiative transfer in the Earth's atmosphere, *BAMS,* **79**, 2101–2114.

Riesz, F., and B. Sz.-Nagy, 1990: *Functional Analysis.* Dover, New York.

Roujean, J.-L., M. Leroy, and P. Y. Deschamps, 1992: A bidirectional reflectance model of the Earth's surface for the correction of the remote sensing data, *J. Geophys. Res.*, **97**, 20,455–20,468.

Schaaf C. B., Gao F., Strahler A.H., W. Lucht, X. Li, T. Tsang, N. C. Strugnell, X. Zhang, Y. Jin, J.-P. Muller, P. Lewis, M. Barnsley, P. Hobson, M. Disney, G. Roberts, M. Dunderdale, C. Doll, R. P. d'Entremont, B. Hu, S. Liang, J. L. Privette and D. Roy, 2002: First operational BRDF, albedo nadir reflectance products from MODIS, *Rem. Sens. Environ.*, **83**, 135–148.

Spurr, R. J. D., Kurosu T. P., Chance K. V., 2001: A linearized discrete ordinate radiative transfer model for atmospheric remote sensing retrieval, *J. Quant. Spectrosc. Radiat. Transfer*, **68**, 689–735.

Stamnes, K., S. C. Tsay, W. Wiscombe and K. Jayaweera, 1988: Numerically stable algorithm for discrete-ordinate-method radiative transfer in multiple scattering and emitting layered media, *Appl. Opt.*, **27**, 2502–2509.

Sushkevich, T. A., S. A. Strelkov, and A. A. Ioltuhovskii, 1990: *Method of Path Integration in the Problems of Atmospheric Optics* (in Russian). Nauka, Moscow, Russia.

Twomey, S., 1985: Green's function formulae for the internal intensity in radiative transfer computations by matrix-vector methods, *J. Quant. Spectrosc. Radiat. Transfer*, **33**, 575–579.

Vladimirov, V. S., 1963: *Mathematical problems in the one-velocity theory of particle transport.* Tech. Rep. AECL-1661, At. Energy of Can. Ltd., Chalk River, Ontario.

Wang, Y., A. Lyapustin, J. L. Privette, J. T. Morisette, B. Holben, 2009: Atmospheric correction at AERONET locations: A new science and validation data set. *IEEE Trans. Geosci. Remote Sens.*, in press.

Wiscombe, W. J., 1977: The Delta-M method: rapid yet accurate radiative flux calculations for strongly asymmetric phase functions. *J. Atm. Sci.*, **34**, 1408–1422.

Zege, E. P., I. L. Katsev, A. P. Ivanov, 1991: *Image Transfer through a Scattering Medium.* Berlin: Springer.

7 General invariance relations reduction method and its applications to solutions of radiative transfer problems for turbid media of various configurations

Nikolai N. Rogovtsov

7.1 Introduction

While solving different theoretical and applied problems of planetary atmosphere optics (Sobolev, 1975; van de Hulst, 1980; Yanovitskij, 1997; Liou, 2002; Kokhanovsky, 2004; Hovenier et al., 2004; Garcia et al., 2008), astrophysics (Mihalas, 1978; Dolginov et al., 1995; Leroy, 2000; Nagirner, 2001), scattering media optics and biophysics (Yi et al., 1992; Tuchin, 1997, 2000; Klose, 2009), signal and image transfer (Zege et al., 1991; Gibson et al., 2005), nuclear plant physics (Davison, 1958; Marchuk, 1961; Marchuk and Lebedev, 1971; Sanchez and McCormick, 1982; Ganapol and Kornreich, 1995; Ganapol, 2008), etc., one should investigate the radiative (neutron) transfer process through macroscopically homogeneous and inhomogeneous, absorbing and scattering media (in particular, through turbid media). The study of the features of this process in various real and artificial absorbing and scattering media requires accounting for their shapes, locations of external and internal sources, kinds of functional dependences of local characteristics of the said media on spatial and angular coordinates. Besides, there should be taken into consideration the physical (optical) and geometrical properties of boundary surfaces and surroundings of absorbing and scattering media. Due to a wide variety of the mentioned characteristics and properties, it is rather difficult to construct unified and effective enough multidimensional (over spatial variables) boundary-value problems of the radiative (neutron) transfer theory. It is most difficult to solve such problems when the shapes of media and the location of sources therein do not possess any constructive symmetry. Among the methods that are universal in many respects, one should note various numerical methods (the description and the evaluation of the opportunities of these methods are given, for example, in the works of Marchuk (1961), Marchuk and Lebedev (1971), Sushkevich et al. (1990), Morel et al. (1996), Adams and Larsen (2002), Santandrea and Sanchez (2002), Sanchez and McCormick (2004), Marshak and Davis (2005), Ganapol and Kornreich (2005) and Nikolaeva et al. (2007)) and the Monte Carlo method (Marchuk et al., 1980; Mikhailov, 1987). Among the analytical (semi-analytical) strict methods enabling one to totally (or at least partially) investigate the properties of the solutions to problems of the radiative (neutron) transfer theory for cases of media and sources having no plane-parallel symmetry, note the method of integral transfor-

mations (Ershov and Shikhov 1972; Kolesov, 1983; Nagirner, 1986, 1994a, 1994b; Ganapol, 2003; Freimanis, 2005), the method of boundary functions (Latishev, 1984), the Case method (Case and Zweifel, 1967; Marchuk and Lebedev, 1971; Kolesov, 1986), and the methods based on the usage of concepts and constructions of functional analysis (Shikhov, 1973; Germogenova, 1986; Germogenova and Pavelyeva, 1991, 1994; Pavelyeva, 1990). Besides, a number of general useful strict expressions and equations have been obtained in the radiative (neutron) transfer theory for macroscopically homogeneous and inhomogeneous media by means of the classical version of the invariant embedding method (Bellman, 1986; Yanovitskij, 1997). It should be noted still that in the radiation transfer theory (RTT) and in the neutron transport theory different combined methods are used rather widely. These methods allow one to take into account a specific character of problem statement and simultaneously to use advantages of the methods mentioned above (see, for example, the following publications: Sobolev, 1975; Ganapol, 1976; Rogovtsov and Samson, 1976; Larsen et al., 1980; Rogovtsov, 1980, 1981a; McCormick and Sanchez, 1981; Sanchez and Ganapol, 1983; Garcia and Siewert, 1983; Adomian et al., 1988; Ganapol and Myneni, 1992; Ganapol et al., 1994; Ganapol and Pomraning, 1996; Yanovitskij, 1997; Ganapol and McCormick, 1997; Rogovtsov, 1999; Kascas et al., 2000; Anisimov and Rogovtsov, 2002; Ganapol and Kornreich, 2005; Cassell and Williams, 2006, 2007; Ganapol, 2008; Rogovtsov and Borovik, 2009; and references therein. Together with numerical methods, Monte Carlo methods, strict methods, etc., the radiative transfer theory also widely uses rather simple approximate analytical methods (see references, for example, in the publications Zege et al. (1991), Tuchin (1997, 2000) and Gorodnichev et al. (2006)). The accuracy of these approximate analytical methods, however, is not high enough always, owing to a number of *a priori* assumptions utilized in their construction, and cannot be evaluated within the scope of the methods themselves.

In addition to the above-said strict (analytical and semi-analytical) methods for solving direct (in particular, multidimensional over spatial variables) and inverse problems of the RTT, one can use the general invariance relations reduction method (GIRRM) that have been proposed by Rogovtsov and Samson (1976); Rogovtsov (1980, 1981a, 1981b, 1981c, 1983, 1985a, 1985b, 1985d, 1989) and presented in the most general form in the monograph (Rogovtsov, 1999). A number of particular, but essentially close results have been obtained in the papers (Pikichian, 1982a, 1982b, 1983, 1989). The general invariance principle (Rogovtsov, 1981a, 1981b, 1999) was the foundation for the general invariance relations reduction method. The important particular consequences of the general invariance principle are the invariance principles by Ambartsumian (1943a, 1943b) and Chandrasekhar (1950), the classical version of the invariant embedding principle (Bellman and Kalaba, 1956), and various generalizations of the said principles that have been stated in the nineteen-sixties to the ninteen-eighties in a number of publications (see, for example, the references in the monographs Yanovitskij (1997) and Rogovtsov (1999)). The distinctive features of the general invariance relations reduction method are, in particular, its heuristicity, constructiveness, and opportunity to get, on its basis, analytical (semi-analytical) solutions to boundary-value problems for the radiative transfer equation (RTE) for the case of an arbitrary phase function and turbid media of complex configurations. With the help of this method, analytical (in par-

ticular asymptotic) or numerical solutions of a number of complex multidimensional boundary-value problems of RTT can be obtained in fact without direct solving of these problems. In essence the general invariance relations reduction method allows one to reduce solving initial boundary-value problems of RTT to the analysis of essentially simpler problems. It should be noted that the construction of GIRRM had been realized using not only general mathematical and physical notions but invoking model considerations of ray optics and probabilistic interpretation of RTT.

The objective of this chapter is the description of main statements of the general invariance relations reduction method and the illustration of its opportunities by means of examples of solving rather complicated multi-dimensional boundary-value problems of the RTT. The layout of this chapter is as follows. Section 7.2 presents briefly the essence of the general invariance relations reduction method. There are given the statements of boundary-value problems for the scalar radiative transfer equation in subsection 7.2.1 for the cases of stationary and non-stationary sources. The general invariance principle (GIP) is formulated in subsection 7.2.2 as applied to the radiative transfer theory. In subsection 7.2.3, a number of the general invariance relations (GIRs), which are the direct consequences of the GIP, are derived. These relations connect together the solutions to similar or different boundary-value problems for the radiative transfer equation. Subsection 7.2.4 gives briefly the description of the scheme, how one can apply the GIP and the GIRs to solve the said problems. A number of general examples illustrating the opportunities of the general invariance relations reduction method are presented in section 7.3 for reducing the original boundary-value problems to the solution to essentially simpler boundary-value problems. Moreover, the analog of the Kirchhoff law for the case of non-equilibrium radiation is obtained in this section and, on the basis of the general invariance relations for the monochromatic fluxes, the strict lower and upper estimations for these fluxes and for mean emission durations of turbid media of non-concave shapes are found. There are derived a number of explicit strict asymptotical and approximate analytical solutions to specific boundary-value problems for the scalar radiative transfer equation in section 7.4 by means of the above-said method (GIRRM). In subsection 7.4.1, the strict expression for the azimuth-averaged reflection function for the case of an semi-infinite plane-parallel non-conservatively scattering turbid medium (its phase function can describe isotropic scattering or be forward extended) is found. There are derived a number of analytical expressions for monochromatic radiation fluxes and luminosity values for turbid media of various shapes in subsection 7.4.2. The opportunities providing by the general invariance relations reduction method to analytically describe depth regimes of radiation fields and to derive mean emission durations of optically thick turbid media are briefly given in subsection 7.4.3.

The main reasons providing sufficient generality, the heuristicity, and the effectiveness of the general invariance relations reduction method are pointed out in the Conclusion (section 7.5). Besides, there are noted the application areas of this method and possible directions of its development. The Appendix contains a brief description of general mathematical notations, conceptions, and constructions, without which it is practically impossible to state strictly the GIP and derive the GIRs. Note that learning the contents of the Appendix can greatly simplify the reader's understanding of the essence of the GIP, the GIRs, and the GIRRM.

7.2 Main statements of the general invariance relations reduction method

7.2.1 Statement of boundary-value problems of the scalar radiative transfer theory

Consider a turbid medium, which occupies a part of the space and does not change its physical (optical) and geometrical characteristics with time (except for characteristics of radiation sources that can, generally, depend on time). Let such a medium be denoted by V_β, where β belongs to a subscript set B. Subscript β determines symbolically the location, surrounding, internal and external sources, as well as physical (optical), geometrical and other properties of V_β. We will also use below the terms of turbid bodies and objects (or simply bodies and objects) instead of the term of a turbid medium. Note here that, in the following, all the objects that are identical each other in geometrical and physical senses at all corresponding points will be always assumed equal and be denoted by the same symbol of the V_β kind. Turbid bodies of various shapes (they can be multi-connected, be disconnected, or be a part of other bodies) can represent V_β. Let S_β be a boundary of a turbid body V_β, the boundary being assumed to possess physical (optical) and geometrical properties. Note that generally any portion of boundary S_β cannot belong to body V_β (i.e. $[V_\beta] = V_\beta$ is not valid always, where $[V_\beta]$ is the closure of set V_β within a topology based on the conception of a distance between points in the three-dimensional Euclidean point space E^3; see, for example, Faure et al. (1964)). Let the conditions that will be imposed on geometrical properties of boundary S_β be specified. Assume that S_β can be represented as a union of a finite or countable set of simple pieces of surfaces, for which a conception of a side and a normal can be introduced at all their internal points (the normals cannot exist at the boundaries of these pieces). In such a case, the direction of an external normal at a specific point of S_β for a selected side of its simple piece should agree with the transition operation through the surface to the other side of the said simple piece of boundary S_β. Let V_β^0 be an internal part of body V_β (i.e. $V_\beta^0 = V_\beta \backslash S_\beta$, where the symbol '$\backslash$' determines the difference operation of sets V_β and S_β). Denote a set of all sides of all simple pieces of the surfaces (their union is S_β), which osculate with V_β^0, via $S_\beta(V_\beta^0)$. Symbols $\vec{n} = \vec{n}(\vec{r}_{S_\beta})$ and $\vec{n}' = \vec{n}'(\vec{r}'_{S_\beta})$ will everywhere below stand for the unit external normals to the selected sides of simple pieces of boundary S_β at points, specified by radius-vectors $\vec{r}_{S_\beta}$ and $\vec{r}'_{S_\beta}$, respectively (they determine S_β in some specified right-handed Cartesian rectangular coordinate system $OXYZ$).

Real and model turbid media can contain underlying surfaces at the external boundary or inside themselves. For such surfaces, local (i.e. determined in certain neighborhoods of points lying on the surfaces) reflection and transmission operators, $\widehat{R}$ and $\widehat{T}$, cannot be simultaneously zero (i.e. in these neighborhoods, the surface does not reflect radiation at all) and unity (i.e. in these neighborhoods, the surface transmits radiation without changing its characteristics), respectively. The local operators can depend on the selection of a side on simple pieces of an underlying surface. From a geometrical viewpoint, some of the underlying surfaces may not be included in the boundary S_β of body V_β. We will assume, however, that these underlying surfaces possess the above-said geometrical properties of the boundary

S_β. Note also that the occurrence of the underlying surfaces is related with the presence of discontinuities of optical characteristics inside the turbid media or at their boundaries.

Let now the scalar radiative transfer equation (SRTE) and boundary-value problems for it be stated under the assumptions that photon frequency does not change (monochromatic scattering) and that local optical characteristics inside turbid medium V_β are isotropic at points, which do not lie on underlying surfaces. Besides, it will be assumed that the local optical characteristics at the said points of body V_β do not depend on the radiation field.

7.2.1.1 The case of non-stationary internal and external sources

Let the above assumptions be valid. Assume that refraction index $m(\vec{r})$ in turbid medium V_β outside of boundary S_β and of underlying surfaces is a differentiable, rather slow changing function of argument $\vec{r}$ (radius-vector $\vec{r}$ specifies an observation point of V_β^0 in some selected right-handed rectangular Cartesian coordinate system; the 'observation' point should not lie on the underlying surfaces). Besides we believe that the inequality of $\left|\left(d\vec{\Omega}/d\tau\right)\cdot\nabla_{\vec{\Omega}} I\left(\vec{r},\vec{\Omega},t;V_\beta\right)\right| \ll I\left(\vec{r},\vec{\Omega},t;V_\beta\right)$ is valid everywhere in V_β^0. Here $\vec{\Omega}$ is the unit vector (i.e. $|\vec{\Omega}|=1$) that specifies the propagation direction of a light ray (this vector defines the direction, corresponding to the ray propagation direction, of a tangent line to the curved trajectory of the light ray in a point set by the end of the radius-vector $\vec{r}$); $I(\vec{r},\vec{\Omega},t;V_\beta)$ is the radiation intensity in turbid medium V_β at time moment t; $d\tau=\kappa_{\text{ext}}(\vec{r})\,dl$, where $\kappa_{\text{ext}}(\vec{r})$ is the extinction coefficient and dl is the length of an infinitesimal element along curved trajectory of the light ray ($d\tau$ has the sense of the optical length of this element); $\nabla_{\vec{\Omega}}$ is the Hamiltonian operator acting only on the variable $\vec{\Omega}$. This inequality imposes essentially the restraint on the curvature of a light ray at any observation point. The curvature should be small enough in regions near boundary S_β near underlying surfaces, and close to the positions of highly anisotropic radiation sources (i.e. changes in the refraction index in such regions should be small enough; however, the refractive index can substantially change, while one goes from one side of a simple piece of boundary S_β (or of an underlying surface) to another its side). Within the scope of the made assumptions, the SRTE obtained by Bekefi (1966) and by Minin (1988) takes the following form:

$$\begin{aligned}
&\frac{1}{v\left(\vec{r}\right)}\frac{\partial I\left(\vec{r},\vec{\Omega},t;V_\beta\right)}{\partial t}+m^2\left(\vec{r}\right)\left(\vec{\Omega}\cdot\nabla_{\vec{r}}\right)\left(\left(m\left(\vec{r}\right)^{-2}\right)I\left(\vec{r},\vec{\Omega},t;V_\beta\right)\right)\\
&=-\kappa_{ext}\left(\vec{r}\right)I\left(\vec{r},\vec{\Omega},t;V_\beta\right)+\frac{\kappa_{sca}\left(\vec{r}\right)}{4\pi}\int\limits_{\Omega}p\left(\vec{r};\,\vec{\Omega}\cdot\vec{\Omega}'\right)I\left(\vec{r},\vec{\Omega},t;V_\beta\right)d\Omega'\\
&\quad+g\left(\vec{r},\vec{\Omega},t;V_\beta^0\right),\quad t\in(-\infty,\,+\infty).
\end{aligned}\tag{7.1}$$

Here, $\nabla_{\vec{r}}$ is the Hamiltonian operator acting only on the variable $\vec{r}$; $v\left(\vec{r}\right)$ is the light velocity in the turbid medium V_β; $\kappa_{sca}\left(\vec{r}\right)$ is the scattering coefficient; Ω is the unit sphere; $\vec{\Omega}'$ is the unit vector $\left(|\vec{\Omega}|=|\vec{\Omega}'|=1\right)$; $p\left(\vec{r};\;\mu\right)$ is the phase function,

which is nonnegative and which for all $\vec{\Omega} \in \Omega$ and for all admissible $\vec{r}$ satisfies the normalization condition $(4\pi)^{-1} \int_{\Omega} p(\vec{r}; \vec{\Omega} \cdot \vec{\Omega}')\, d\Omega' = 1$ $(\vec{\Omega} \cdot \vec{\Omega}' = \mu \in [-1, 1]\,;\ \vec{\Omega} \cdot \vec{\Omega}'$ is the scalar product of vectors $\vec{\Omega}$ and $\vec{\Omega}'$); symbol $\vec{\Omega} \cdot \nabla_{\vec{r}}$ means the differentiation operation with respect to the direction of unit vector $\vec{\Omega}$; function $g(\vec{r}, \vec{\Omega}, t; \mathrm{V}_\beta^0)$ specifies the density of internal sources in V_β^0. In Eq. (7.1), function $I(\vec{r}, \vec{\Omega}, t; \mathrm{V}_\beta)$ is unknown and all other functions are supposed to be specified. Generally, all the functions entering Eq. (7.1) can be dependent on frequency ν as a parameter.

The strict mathematical statement of boundary-value problems for Eq. (7.1) not only requires one to form the boundary conditions imposed on boundary S_β of body V_β and of underlying surfaces, but also supposes the indication of all function classes entering this equation. With this, the selected function classes should agree with each other, so that the solution to Eq. (7.1) will exist in the respective function class. As with any other equation comprising differential operations, Eq. (7.1) imposes only general restrictions on the function class, where the solution $I(\vec{r}, \vec{\Omega}, t; \mathrm{V}_\beta)$ is sought. To isolate required and physically reasonable solutions, a number of additional conditions should be added to Eq. (7.1).

While stating a general boundary-value problem for Eq. (7.1), one should take into account that external sources and surrounding of turbid medium V_β affect the radiation field in the medium (in particular, V_β may be a part of another turbid medium). Besides, radiation incident on $\mathrm{S}_\beta(\mathrm{V}_\beta^0)$ (or on underlying surfaces, if they are in V_β) from the interior of body V_β can again strike the same turbid medium. Radiation exiting from the interior of such a medium (i.e. from V_β^0) can, for example, again partially or totally come to V_β^0 as a result of reflection, transmission, refraction, or 'shooting' ('shooting' means mutual irradiation of the same object through spatial regions not belonging to V_β). Therefore, radiation exiting V_β^0 near a point specified by $\vec{r}_s$ does not necessarily come to V_β^0 again from the neighborhood of the same point.

Accounting for the above, consider a turbid medium V_β, the boundary S_β of which includes all underlying surfaces of V_β. Introduce hemispheres $\Omega_\pm = \Omega_\pm(\vec{r}_s)$ of a unit sphere Ω, which are determined by conditions $\vec{n}\,(\vec{r}_s) \cdot \vec{\Omega} > 0$ and $\vec{n}\,(\vec{r}_s) \cdot \vec{\Omega} < 0$, respectively (vector $\vec{n}$ in these inequalities can be replaced by $\vec{n}' = \vec{n}'(\vec{r}'_{\mathrm{S}_\beta})$ with the simultaneous replacement of $\vec{r}_{\mathrm{S}_\beta}$ by $\vec{r}'_{\mathrm{S}_\beta}$). General boundary conditions can be stated, if spatial region $E^3 \backslash \mathrm{V}_\beta^0$ possesses so properties that, by the intensity specified on $\mathrm{S}_\beta(\mathrm{V}_\beta^0)$ for hemispheres of the Ω_+ kind, one can unambiguously find the radiation intensity on $\mathrm{S}_\beta(\mathrm{V}_\beta^0)$ for hemispheres of the Ω_- kind. The above considerations within the scope of the assumptions on the independence of optical and other properties (except for the properties of internal and external sources) of body V_β and its surrounding from the radiation field and time enable one to write down the following boundary condition for function $I(\vec{r}, \vec{\Omega}, t; \mathrm{V}_\beta)$:

$$I(\vec{r}_{\mathrm{S}_\beta} - 0\vec{n}, \vec{\Omega}, t; \mathrm{V}_\beta) = I_{ext}(\vec{r}_{\mathrm{S}_\beta} - 0\vec{n}, \vec{\Omega}, t; \mathrm{V}_\beta) + \iint\limits_{\mathrm{S}_\beta(\mathrm{V}_\beta^0)} dS' \int\limits_{\Omega_+} d\Omega' \int\limits_{-\infty}^{t} R_*(\vec{r}_{\mathrm{S}_\beta} - 0\vec{n}; \vec{\Omega}; \vec{r}', \vec{\Omega}'; t - t'; \mathrm{V}_\beta) I(\vec{r}', \overrightarrow{\Omega'}, t'; \mathrm{V}_\beta)\, dt'. \tag{7.2}$$

Here $\vec{\Omega} \in \Omega_-\,(\vec{r}_S) = \Omega_-;\ t \in (-\infty, +\infty);\ \vec{r}_{\mathrm{S}_\beta} - 0\vec{n}$ is the radius-vector defining a point lying on $\mathrm{S}_\beta(\mathrm{V}_\beta^0)$ (i.e. on the infinitesimal distance from S_β); $I_{ext}(\dots)$ is the specified intensity of external radiation transmitted through the boundary S_β; $R_*(\dots)$ is the known function characterizing optical and geometrical properties of the boundary S_β and of the surrounding of body V_β (the surrounding is in $E^3 \backslash \mathrm{V}_\beta^0$). The surface integral of Eq. (7.2) is being taken over all the sides of simple pieces of surfaces, the union of which is S_β and which osculate with V_β^0. The independence of the characteristics of body V_β and its surrounding from time and the homogeneity of time itself give rise to the fact that function $R_*(\dots)$ should depend on the difference $(t - t')$ only (this fact is just included in Eq. (7.2)). Partial cases of Eq. (7.2) are boundary conditions specified by Germogenova (1986); Rogovtsov (1999).

Since Eq. (7.1) comprises the partial derivative of the first order with respect to t, it is also necessary to set the initial condition in V_β^0

$$I\big(\vec{r}, \vec{\Omega}, t; \mathrm{V}_\beta\big)|_{t=t_0-0} = a\big(\vec{r},\ \vec{\Omega}\big), \tag{7.3}$$

where $t_0 \in (-\infty, +\infty)$ and $a\big(\vec{r},\ \vec{\Omega}\big)$ is the specified function.

In the radiative transfer theory, the first boundary-value problem plays an important role. The problem stands for a problem of solving Eq. (7.1) with accounting for the boundary condition of Eq. (7.2) at $R_*(\dots) \equiv 0$ and the initial condition of Eq. (7.2). The second boundary-value problem will be a problem for solving Eq. (7.1) with accounting for the initial condition of Eq. (7.3) and the boundary condition of Eq. (7.2), when function $I\big(\vec{r}_{\mathrm{S}_\beta}, \vec{\Omega}, t; \mathrm{V}_\beta\big)$ on $\mathrm{S}_\beta(\mathrm{V}_\beta^0)$ is assumed to be specified for all $t \in \big(-\infty,\ t_0\big)$ and $\vec{\Omega} \in \Omega_+\,\big(\vec{r}_{\mathrm{S}_\beta}\big)$, and $I_{ext}(\dots) \equiv 0$ for all $t \in (-\infty, +\infty)$. The problem of Eqs (7.1)–(7.3) without the above restraints is called as a general boundary-value problem. Basic results concerning the qualitative mathematical investigations of boundary-value problems and other problems of the radiative (neutron) transfer theory were obtained, in particular, by Maslennikov (1968); Vladimirov (1971); Marchuk and Lebedev (1971); Shikhov (1973); Germogenova and Shulaya (1976); Germogenova (1986); Germogenova and Pavelyeva (1989); Rogovtsov and Borovik (2009). Although these results will not be described explicitly in the following, some of them will be accounted for, while stating the general invariance relations reduction method (GIRRM) and while deriving the solutions to specific boundary-value problems for Eq. (7.1).

As for a lot of topics of mathematical physics, the Green functions for Eq. (7.1) play an important role in the linear radiative transfer theory. Let us define these functions. Volume Green function $G_*(\vec{r},\ \vec{\Omega},\ \vec{r}', \vec{\Omega}', t; \mathrm{V}_\beta)$ for Eq. (7.1) will be stood for a solution to the second boundary-value problem under the following assumptions: (i) $I\big(\vec{r}, \vec{\Omega}, t; \mathrm{V}_\beta\big)|_{t<0} \equiv, 0$ for all points in V_β^0 and on $\mathrm{S}_\beta(\mathrm{V}_\beta^0)$ as well as for all $\vec{\Omega} \in \Omega$; (ii) $g\big(\vec{r}, \vec{\Omega}, t; \mathrm{V}_\beta^0\big) = \delta\big(\vec{r} - \vec{r}'\big)\delta(\vec{\Omega} - \vec{\Omega}')\,\delta(t)$, where $\vec{r}'$ specifies a point in V_β^0, and $\delta(\dots)$ is the Dirac function (see, for example, Case and Zweifel (1967) and Vladimirov (1979). Note that function $E_{00}G_*(\vec{r},\ \vec{\Omega},\ \vec{r}', \vec{\Omega}', t; \mathrm{V}_\beta)$ determines the radiation intensity in body V_β, when there is a point monodirectional radiation source inside the body, which emits instantaneously energy E_{00} at time moment $t = 0$ within unit spectral range. Surface Green function

$G_{\mathrm{S}_\beta}(\vec{r},\ \vec{\Omega},\ \vec{r}\,'_{S_\beta}-0\vec{n}',\vec{\Omega}',t;\mathrm{V}_\beta)$ for Eq. (7.1) is the solution to the general boundary-value problem of Eqs (7.1)–(7.3), when the following conditions are fulfilled: (*i*) $I\big(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta\big)|_{t<0}\equiv 0$ for all point in V^0_β and on $\mathrm{S}_\beta(\mathrm{V}^0_\beta)$ as well as for all $\vec{\Omega}\in\Omega$; (*ii*) $g\big(\vec{r},\vec{\Omega},t;\mathrm{V}^0_\beta\big)|_{t\in(-\infty,+\infty)}\equiv 0$ for all points in V^0_β and for all $\vec{\Omega}\in\Omega$; (*iii*) $I_{ext}\big(\vec{r}_{\mathrm{S}_\beta}-0\vec{n},\vec{\Omega},t;\mathrm{V}_\beta\big)=\delta_{\mathrm{S}_\beta}\big(\vec{r}_{\mathrm{S}_\beta}-\vec{r}\,'_{\mathrm{S}_\beta}\big)\delta(\vec{\Omega}-\vec{\Omega}')\,\delta(t)$, where $\vec{n}=\vec{n}\,(\vec{r}_s)$, $\vec{n}'\cdot\vec{\Omega}'<0$ $\big(\vec{n}'=\vec{n}'\big(\vec{r}\,'_{S_\beta}\big)\big)$, and $\delta_{\mathrm{S}_\beta}(\ldots)$ is the surface delta function (Case and Zweifel (1967) and Vladimirov (1971)). Function $I_0G_{\mathrm{S}_\beta}(\vec{r},\ \vec{\Omega},\ \vec{r}\,'_{\mathrm{S}_\beta}-0\vec{n}',\vec{\Omega}',t;\mathrm{V}_\beta)$ specifies radiation intensity in body V_β, when a surface mono-directional instantaneous radiation source with surface density $I_0\big|\vec{n}'\big(\vec{r}\,'_{\mathrm{S}_\beta}\big)\cdot\vec{\Omega}'\big|\delta_{\mathrm{S}_\beta}\big(\vec{r}_{\mathrm{S}_\beta}-\vec{r}\,'_{\mathrm{S}_\beta}\big)\delta(\vec{\Omega}-\vec{\Omega}')\,\delta(t)$ is one the internal side of a simple piece of the body boundary (i.e. on $\mathrm{S}_\beta(\mathrm{V}^0_\beta)$).

If turbid medium V_β is irradiated by internal and external sources, is disconnected or is embedded in another medium (in particular, in a turbid medium), and contains underlying surfaces, then, generally, the derivation of functions $I_{ext}(\ldots)$ and $R_*(\ldots)$ entering the boundary condition of Eq. (7.2) and the solution to boundary-value problems for Eq. (7.1) are rather problematic. For such situations, the radiative transfer should be treated simultaneously in turbid medium V_β itself and in its surrounding. Then the boundary condition of Eq. (7.2) takes the simplest form, when function $R_*(\ldots)$ with respect to $(t-t')$ is proportional to $\delta\,(t-t')$.

Let body V_β be surrounded by a medium that is able to only absorb (or not absorb) radiation (but not to scatter it) and to contain specified external sources. If optical properties of all underlying surfaces included in V_β and characteristics of external sources are known, then, in principle, one can find explicit expressions for functions $I_{ext}(\ldots)$ and $R_*(\ldots)$. Besides, the solution to the general boundary-value problem of Eqs (7.1)–(7.3) in this case can be written down in the following form:

$$
\begin{aligned}
&I\big(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta\big)\\
&=\iint\limits_{\mathrm{S}_\beta(\mathrm{V}^0_\beta)}dS'\int\limits_{\Omega_-}d\Omega'\int\limits_{-\infty}^{t}G_{\mathrm{S}_\beta}\big(\vec{r};\vec{\Omega};\vec{r}\,',\vec{\Omega}';t-t';\mathrm{V}_\beta\big)I_{ext}\big(\vec{r}\,',\overrightarrow{\Omega'},t';\mathrm{V}_\beta\big)\,dt'\\
&\quad+\iiint\limits_{\mathrm{V}^0_\beta}d\mathrm{V}'\int\limits_{\Omega}d\Omega'\int\limits_{-\infty}^{t}G_*\big(\vec{r};\vec{\Omega};\vec{r}\,',\vec{\Omega}';t-t';\mathrm{V}_\beta\big)g\big(\vec{r}\,',\vec{\Omega}';t';\mathrm{V}^0_\beta\big)\,dt',\qquad(7.4)
\end{aligned}
$$

where $t\in(-\infty,+\infty)$, $\vec{\Omega}\in\Omega$. For this, radius vector $\vec{r}$ of Eq. (7.4) specifies points in V^0_β.

In the radiative transfer theory, it is sometimes useful to use equations of the integral form instead of the SRTE. Let body V_β is non-concave, does not contain any underlying surfaces, and is not imbedded in another turbid medium. At this, a non-concave body stands for a body, which, together with any two of its points, contains completely the line segment connecting them. Apply now the Laplace transform to Eq. (7.1). One obtains finally

$$
m^2(\vec{r}\,)\big(\vec{\Omega}\cdot\nabla_{\vec{r}}\big)\Big(\big(m\big(\vec{r}\,\big)\big)^{-2}\bar{I}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\beta\big)\Big)+\kappa^*_{ext}\,(\vec{r},\ \sigma)\,\bar{I}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\beta\big)=\bar{\Phi}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\beta\big).\qquad(7.5)
$$

Here σ is the parameter of the Laplace transform $\bar{f}(\sigma) = \hat{L}(f)(\sigma) = \int_{-\varepsilon}^{+\infty} f(t)\exp(-\sigma t)\,dt$ of function $f(t)$ (it will be assumed below that σ is always a nonnegative number, and ε will stand for an infinitesimal positive number); $\kappa^*_{ext}(\vec{r};\ \sigma) = \kappa_{ext}(\vec{r}) + \sigma\,(v(\vec{r}))^{-1}$;

$$\bar{\Phi}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\beta\big) = \frac{\kappa_{sca}(\vec{r})}{4\pi}\int\limits_{\Omega} p\big(\vec{r};\,\vec{\Omega}\cdot\vec{\Omega}'\big)\,\bar{I}\big(\vec{r},\vec{\Omega}',\sigma;\mathrm{V}_\beta\big)\,d\Omega' + \bar{g}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}^0_\beta\big) + \big(v(\vec{r})\big)^{-1} I\big(\vec{r},\vec{\Omega},-0;\mathrm{V}_\beta\big).$$

If function $\bar{I}_{ext}\big(\vec{r}_{\mathrm{S}_\beta} - 0\vec{n}\,\big(\vec{r}_{\mathrm{S}_\beta}\big),\vec{\Omega},\sigma;\mathrm{V}_\beta\big)$ is assumed to be known for all $\vec{r}_{\mathrm{S}_\beta}$ and for all $\vec{\Omega}\in\Omega_-\big(\vec{r}_{\mathrm{S}_\beta}\big)$, then, using Eq. (7.5) and the definition of function $\bar{\Phi}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\beta\big)$, one can get the following equation of the integral form (Rogovtsov, 1981a, 1999):

$$\begin{aligned}\bar{\Phi}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\beta\big) &= \big(v(\vec{r})\big)^{-1} I\big(\vec{r},\vec{\Omega},-0;\mathrm{V}_\beta\big) + \bar{g}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}^0_\beta\big)\\ &+ (4\pi)^{-1}\kappa_{sca}(\vec{r}) \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right)} m_1\big(\vec{r},\vec{r}\,'\big)p\big(\vec{r};\vec{\Omega}\cdot\vec{\rho}\big)\bar{I}_{ext}\big(\vec{r}\,',\vec{\rho},\sigma;\mathrm{V}_\beta\big)\\ &\cdot \exp\big(-\eta(\vec{r},\vec{r}\,',\sigma)\big)\,|\vec{n}'\cdot\vec{\rho}|\,\frac{dS'}{\rho_1^2}\\ &+ (4\pi)^{-1}\kappa_{sca}(\vec{r}) \iiint\limits_{\mathrm{V}^0_\beta} m_1\big(\vec{r},\vec{r}\,'\big)p\big(\vec{r};\vec{\Omega}\cdot\vec{\rho}\big)\\ &\cdot \exp\big(-\eta(\vec{r},\vec{r}\,',\sigma)\big)\,\bar{\Phi}(\vec{r}\,',\vec{\rho},\sigma;\mathrm{V}_\beta)\,\frac{d\mathrm{V}'}{\rho_1^2}. \end{aligned} \tag{7.6}$$

Here radius-vector $\vec{r}$ specifies points in V^0_β; $\vec{\Omega}\in\Omega$, $\vec{\rho} = (\vec{r}-\vec{r}\,')\,\rho_1^{-1}$, $\rho_1 = |\vec{r}-\vec{r}\,'|$; $\eta(\vec{r},\vec{r}\,',\sigma) = \int_0^{\rho_1}\kappa^*_{ext}(\vec{r}-\xi\vec{\rho},\sigma)\,d\xi$, $\vec{n}' = \vec{n}'\big(r'_\beta\big)$; $m_1(\vec{r},\vec{r}\,') = \big(m(\vec{r})/m(\vec{r}\,')\big)^2$. Functions $I\big(\vec{r},\vec{\Omega},-0;\mathrm{V}_\beta\big)$ and $\bar{g}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}^0_\beta\big)$ are assumed in Eq. (7.6) to be known for any points of V^0_β and for all $\vec{\Omega}\in\Omega$. If Eq. (7.6) is solved with respect to function $\bar{\Phi}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\beta\big)$, then, using Eq. (7.5) and the definition of this function, one can derive the following expression for function $\bar{I}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\beta\big)$:

$$\begin{aligned}&\bar{I}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\beta\big)\\ &= m_1\big(\vec{r},\vec{r}-\xi_0\vec{\Omega}\big)\bar{I}_{ext}\big(\vec{r}-\xi_0\vec{\Omega},\vec{\Omega},\sigma;\mathrm{V}_\beta\big)\,\exp\left(-\int\limits_0^{\xi_0}\kappa^*_{ext}(\vec{r}-\xi'\vec{\Omega},\sigma)d\xi'\right) +\\ &+\int\limits_0^{\xi_0} m_1\big(\vec{r},\vec{r}-\xi'\vec{\Omega}\big)\exp\left(-\int\limits_0^{\xi'}\kappa^*_{ext}(\vec{r}-\xi''\vec{\Omega},\sigma)d\xi'\right)\bar{\Phi}(\vec{r}-\xi'\vec{\Omega},\vec{\Omega},\sigma;\mathrm{V}_\beta)d\xi. \end{aligned} \tag{7.7}$$

Here radius-vector $\vec{r}$ specifies points in V^0_β; $\vec{\Omega}\in\Omega$, $\big(\vec{r}-\xi_0\vec{\Omega}\big)$ is the radius-vector of the intersection point of the straight line passing through the end of radius vector $\vec{r}$ to the direction of vector $(-\vec{\Omega})$ with surface S_β.

If in Eqs (7.6) and (7.7) one comes from Laplacian images $\bar{\Phi}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\beta\big)$ and $\bar{I}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\beta\big)$ to their originals $\Phi\big(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta\big)$ and $I\big(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta\big)$ we get (Rogovtsov, 1981a, 1999) the following:

$$
\begin{aligned}
&\Phi\big(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta\big)\\
&=\ (v\,(\vec{r}))^{-1}\,I\big(\vec{r},\vec{\Omega},-0;\mathrm{V}_\beta\big)\delta(t)+g\big(\vec{r},\vec{\Omega},t;\mathrm{V}^0_\beta\big)\\
&\quad+\ (4\pi)^{-1}\,\kappa_{sca}\,(\vec{r})\iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right)} m_1\,(\vec{r},\vec{r}\,')p\big(\vec{r};\vec{\Omega}\cdot\vec{\rho}\big)\exp\left(-\int\limits_0^{\rho_1}\kappa_{ext}(\vec{r}-\xi\vec{\rho})d\xi\right)\\
&\quad\times\,\Theta\left(t-\int\limits_0^{\rho_1}(v(\vec{r}-\xi\vec{\rho}))^{-1}d\xi+\varepsilon\right)\\
&\quad\times\,I_{ext}\left(\vec{r}\,',\vec{\rho},t-\int\limits_0^{\rho_1}(v(\vec{r}-\xi\vec{\rho}))^{-1}d\xi;\mathrm{V}_\beta\right)|\vec{n}\,'\cdot\vec{\rho}|\,\frac{dS'}{\rho_1^2}\\
&\quad+\ (4\pi)^{-1}\,\kappa_{sca}\,(\vec{r})\iiint\limits_{\mathrm{V}^0_\beta} m_1\,(\vec{r},\vec{r}\,')p\big(\vec{r};\vec{\Omega}\cdot\vec{\rho}\big)\exp\left(-\int\limits_0^{\rho_1}\kappa_{ext}(\vec{r}-\xi\vec{\rho})d\xi\right)\\
&\quad\times\,\Theta\left(t-t_1\left(\vec{r},\vec{\rho};\rho_1\right)+\varepsilon\right)\Phi\left(\vec{r}\,',\vec{\rho},t-t_1\left(\vec{r},\vec{\rho};\rho_1\right);\mathrm{V}_\beta\right)\frac{d\mathrm{V}'}{\rho_1^2};
\end{aligned}
\tag{7.8}
$$

$$
\begin{aligned}
&I\left(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta\right)\\
&=\ m_1\left(\vec{r},\vec{r}-\xi_0\vec{\Omega}\right)\exp\left(-\int\limits_0^{\xi_0}\kappa_{ext}(\vec{r}-\xi'\vec{\Omega})d\xi'\right)\Theta\left(t-t_1\left(\vec{r},\vec{\Omega};\ \xi_0\right)+\varepsilon\right)\\
&\quad\times\,I_{ext}\big(\vec{r}-\xi_0\vec{\Omega},\vec{\Omega},t-t_1\big(\vec{r},\vec{\Omega};\ \xi_0\big);\mathrm{V}_\beta\big)\\
&\quad+\int\limits_0^{\xi_0} m_1\left(\vec{r},\vec{r}-\xi'\vec{\Omega}\right)\exp\left(-\int\limits_0^{\xi'}\kappa_{ext}(\vec{r}-\xi''\vec{\Omega})d\xi''\right)\\
&\quad\times\,\Theta\big(t-t_1\big(\vec{r},\vec{\Omega};\ \xi'\big)+\varepsilon\big)\Phi\big(\vec{r}-\xi'\vec{\Omega},\vec{\Omega},t-t_1\big(\vec{r},\vec{\Omega};\ \xi'\big);\mathrm{V}_\beta\big)\,d\xi'.
\end{aligned}
\tag{7.9}
$$

Here $t_1\big(\vec{r},\vec{b};u\big)=\int_0^u\left(v\big(\vec{r}-\xi\vec{b}\big)\right)^{-1}d\xi$, where $\big|\vec{b}\big|=1$; $\Theta\,(x)=1$ for any $x\in[0,\ +\infty)$ and $\Theta\,(x)=0$ for any $x<0$.

7.2.1.2 The case of stationary internal and external sources

If external and internal sources are independent of time t, the investigation of radiative transfer through turbid media of various configurations becomes considerably simpler. State the stationary analogs of the equations, definitions, and relations

that have been given above for the case of non-stationary sources. At this, let all the above-made assumptions not concerning the radiation sources be valid.

The stationary analog of Eq. (7.1) is the following:

$$m^2(\vec{r})\left(\vec{\Omega}\cdot\nabla_{\vec{r}}\right)\left(m(\vec{r})^{-2}I\left(\vec{r},\vec{\Omega};\mathrm{V}_\beta\right)\right) = -\kappa_{ext}(\vec{r})\,I\left(\vec{r},\vec{\Omega};\mathrm{V}_\beta\right) + (4\pi)^{-1}\kappa_{sca}(\vec{r})\int_\Omega p\left(\vec{r},\vec{\Omega}\cdot\vec{\Omega}'\right)I\left(\vec{r},\vec{\Omega}';\mathrm{V}_\beta\right)d\Omega' + g\left(\vec{r},\vec{\Omega};\mathrm{V}_\beta^0\right). \quad (7.10)$$

Here $I\left(\vec{r},\vec{\Omega};\mathrm{V}_\beta\right)$ is the radiation intensity in turbid medium V_β^0; $g\left(\vec{r},\vec{\Omega};\mathrm{V}_\beta^0\right)$ is the density of internal stationary sources in V_β^0.

The stationary analog of the general boundary conditions of Eq. (7.2) is

$$I\left(\vec{r}_{\mathrm{S}_\beta}-0\vec{n},\vec{\Omega};\mathrm{V}_\beta\right) = I_{ext}\left(\vec{r}_{\mathrm{S}_\beta}-0\vec{n},\vec{\Omega};\mathrm{V}_\beta\right) + \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)} dS' \int\limits_{\Omega_+} R_*\left(\vec{r}_{\mathrm{S}_\beta}-0\vec{n},\vec{\Omega};\vec{r}',\vec{\Omega}';\mathrm{V}_\beta\right)I\left(\vec{r}';\vec{\Omega}';\mathrm{V}_\beta\right)d\Omega', \quad (7.11)$$

where $R_*\left(\vec{r}_{\mathrm{S}_\beta}-0\vec{n},\vec{\Omega};\vec{r}',\vec{\Omega}';\mathrm{V}_\beta\right) = \int_0^{+\infty} R_*\left(\vec{r}_{\mathrm{S}_\beta}-0\vec{n},\vec{\Omega};\vec{r}',\vec{\Omega}';t;\mathrm{V}_\beta\right)dt$, $\vec{\Omega}\in\Omega_-\left(\vec{r}_S\right)=\Omega_-$, and $I_{ext}\left(\vec{r}_{\mathrm{S}_\beta}-0\vec{n},\vec{\Omega};\mathrm{V}_\beta\right)$ is the radiation intensity generated by the external stationary sources.

The first boundary-value problem for the case of stationary sources stands for a problem of solving Eq. (7.10) with accounting for the boundary condition of Eq. (7.11) at $R_*(\ldots)\equiv 0$ (Germogenova, 1986). The second boundary-value problem is, respectively, a problem of solving Eq. (7.10) with accounting for the boundary condition of Eq. (7.11) at $I_{ext}(\ldots)\equiv 0$ (Germogenova, 1986). The volume Green function $G_*(\vec{r},\ \vec{\Omega};\ \vec{r}',\vec{\Omega}';\mathrm{V}_\beta)$ for Eq. (7.10) is a solution to the second boundary-value problem, when $g\left(\vec{r},\vec{\Omega};\mathrm{V}_\beta^0\right)=\delta(\vec{r}-\vec{r}')\,\delta\left(\vec{\Omega}-\vec{\Omega}'\right)$, $\vec{r}$ and $\vec{r}'$ specify points in V_β^0. The surface Green function $G_{\mathrm{S}_\beta}\left(\vec{r},\vec{\Omega},\vec{r}'_{\mathrm{S}_\beta}-0\vec{n}',\vec{\Omega}';\mathrm{V}_\beta\right)$ for Eq. (7.10) is a solution to the general boundary-value problem, when $g\left(\vec{r},\vec{\Omega};\mathrm{V}_\beta^0\right)\equiv 0$ and $I_{ext}(\vec{r}_{\mathrm{S}_\beta}-0\vec{n},\vec{\Omega};\mathrm{V}_\beta)=\delta_{\mathrm{S}_\beta}\left(\vec{r}_{\mathrm{S}_\beta}-\vec{r}'_{\mathrm{S}_\beta}\right)\delta\left(\vec{\Omega}-\vec{\Omega}'\right)$, where $\vec{n}'\cdot\vec{\Omega}'<0$ $\left(\vec{n}'=\vec{n}'\left(\vec{r}'_{\mathrm{S}_\beta}\right)\right)$.

The stationary analog of Eq. (7.4) is

$$I\left(\vec{r},\vec{\Omega};\mathrm{V}_\beta\right) = \iint\limits_{\mathrm{S}_\beta(\mathrm{V}_\beta^0)} dS' \int\limits_{\Omega_-} G_{\mathrm{S}_\beta}\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}';\mathrm{V}_\beta\right)I_{ext}\left(\vec{r}',\overrightarrow{\Omega'};\mathrm{V}_\beta\right)d\Omega' + \iiint\limits_{\mathrm{V}_\beta^0} d\mathrm{V}' \int\limits_{\Omega} G_*\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}';\mathrm{V}_\beta\right)g\left(\vec{r}',\vec{\Omega}';\mathrm{V}_\beta^0\right)d\Omega', \quad (7.12)$$

where $\vec{\Omega}\in\Omega$, and $\vec{r}$ specifies points in V_β^0.

The stationary analog of Eqs (7.6) and (7.8) is (Rogovtsov, 1981a, 1999)

$$\begin{aligned}\Phi\big(\vec{r},\vec{\Omega};\mathrm{V}_\beta\big) &= g\big(\vec{r},\vec{\Omega},\mathrm{V}_\beta^0\big)\\ &+ (4\pi)^{-1}\kappa_{sca}(\vec{r}) \iint\limits_{\mathrm{S}_\beta(\mathrm{V}_\beta^0)} m_1\big(\vec{r},\vec{r}\,'\big)p\big(\vec{r};\vec{\Omega}\cdot\vec{\rho}\big)I_{ext}(\vec{r}\,';\vec{\rho};\mathrm{V}_\beta)\\ &\times \exp\left(-\eta(\vec{r},\vec{r}\,',0)\right)|\vec{n}'\cdot\vec{\rho}|\,\frac{dS'}{\rho_1^2}\\ &+ (4\pi)^{-1}\kappa_{sca}(\vec{r}) \iiint\limits_{\mathrm{V}_\beta^0} m_1\big(\vec{r},\vec{r}\,'\big)p\big(\vec{r};\vec{\Omega}\cdot\vec{\rho}\big)\\ &\times \exp\left(-\eta(\vec{r},\vec{r}\,',0)\right)\Phi(\vec{r}\,',\vec{\rho};\mathrm{V}_\beta)\,\frac{d\mathrm{V}'}{\rho_1^2}. \end{aligned} \tag{7.13}$$

Here $\vec{r}$ specifies points in V_β^0, $\vec{\Omega}\in\Omega$, and the desired function $\Phi\big(\vec{r},\vec{\Omega};\mathrm{V}_\beta\big)$ is determined by

$$\Phi\big(\vec{r},\vec{\Omega};\mathrm{V}_\beta\big) = (4\pi)^{-1}\kappa_{sca}(\vec{r})\int\limits_{\Omega} p\big(\vec{r},\vec{\Omega}\cdot\vec{\Omega}'\big)I\big(\vec{r},\vec{\Omega}';\mathrm{V}_\beta\big)\,d\Omega' + g\big(\vec{r},\vec{\Omega};\mathrm{V}_\beta^o\big). \tag{7.14}$$

The stationary analog of Eqs (7.7) and (7.9) is

$$\begin{aligned} I\big(\vec{r},\vec{\Omega};\mathrm{V}_\beta\big) &= m_1\big(\vec{r},\vec{r}-\xi_0\vec{\Omega}\big)I_{ext}\big(\vec{r}-\xi_0\vec{\Omega},\vec{\Omega};\mathrm{V}_\beta\big)\,\exp\left(-\int\limits_0^{\xi_0}\kappa_{ext}(\vec{r}-\xi'\vec{\Omega})\,d\xi'\right)\\ &+ \int\limits_0^{\xi_0} m_1\big(\vec{r},\vec{r}-\xi'\vec{\Omega}\big)\,\exp\left(-\int\limits_0^{\xi'}\kappa_{ext}(\vec{r}-\xi''\vec{\Omega})\,d\xi''\right)\Phi(\vec{r}-\xi'\vec{\Omega},\vec{\Omega};\mathrm{V}_\beta)\,d\xi', \end{aligned} \tag{7.15}$$

where $\vec{r}$ specifies points in V_β^0, $\vec{\Omega}\in\Omega$, and number ξ_0 has the same sense as in Eq. (7.7).

7.2.2 Statement of the general invariance principle as applied to radiative transfer theory

Let the problems related with using a conception of the invariance in the radiative transfer theory and with stating the general invariance principle (GIP) be considered now. The GIP was first stated by Rogovtsov (1981a, 1981b) to enable the study of the radiative transfer through turbid media of various configurations from rather general viewpoints. The configuration always stands below for not only a shape of a turbid medium, but also a type of location or distribution of radiation sources inside or outside the medium. The most common treatment of the GIP embracing all the most essential aspects of the invariance (symmetry) principles utilized in physics and mathematics was made by Rogovtsov (1999). The features of the GIP (Rogovtsov, 1981a, 1981b, 1999) are its heuristicity, efficiency, and applicability to the solution to comprehensive boundary-value problems of the radiative (neutron)

transfer theory and to other problems of mathematical physics. Besides, on the basis of the GIP, one succeeds in realizing the relations and differences existing between ways of using the invariance conceptions in mathematics, in theoretical physics, and in radiative transfer theory itself. In particular, it has been strictly proved by Rogovtsov (1981a, 1981b, 1999) that the GIP is not a symmetry principle within the scope of radiative transfer theory (for example, all the invariance principles underlying the foundations of modern physics are symmetry principles).

Theoretical investigation of the radiative transfer through turbid media of complex configurations requires accounting for their shapes, local optical properties of underlying surfaces and of internal portions of these media, types of location and angular patterns of internal and external radiation sources, etc. Due to the integral character of the radiative transfer, radiation fields in turbid media can substantially change with the above characteristics and properties of such media. An enormous number of real and artificial turbid media, which essentially differ from each other, makes rather comprehensive the construction of universal analytical and semi-analytical methods for solving boundary-value problems of the radiative transfer theory. One of such ways is the general invariance relations reduction method (GIRRM) being based on the general invariance principle.

Now a number of evidences, which will be utilized for stating the GIP as applied to the radiative transfer theory, will be given. The definitions and constructions concerning the general treatment of the GIP are presented by Rogovtsov (1999).

Consider a set $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$, the elements of which are arbitrary turbid media (bodies, objects) V_β. Any of the elements of set $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$ can be a part of another body. Let set $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$ include all abstract-realizable objects that can be mentally constructed with various algebraic operations being described below. Assume that the restraints pointed out in subsection 7.2 before the statement of Eq. (7.1) are imposed on all elements of V_β and their boundaries S_β as well as on underlying surfaces. Besides, assume that $\mathrm{V}_{\beta_1} \neq \mathrm{V}_{\beta_2}$, if $\beta_1 \neq \beta_2$. Let a set $\{\Xi\}_{j\in\mathrm{J}}$ of surfaces, each of which has zero measure at arbitrary time moment t in any bounded subset of three-dimensional Euclidean point space E^3, be specified (the measure here and below stands for a measure in this space). In addition to surfaces, curves and finite (countable) sets of points have zero measure in E^3 too. Symbol J is assumed to be a subscript set. Let all the surfaces of the Ξ_j type possess the geometrical properties inherent to boundaries S_β (the properties are described in subsection 7.2). Assume that the positions of the elements of sets $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$ and $\{\Xi_j\}_{j\in\mathrm{J}}$ are being specified with respect to the same coordinate system related unambiguously with system *OXYZ*. It should be noted that really existing turbid media and artificially made objects are always completely separated in the space or have physically interacting common parts between them (e.g. a body may be a part of another body). But mentally created objects in the same coordinate system can be, in an abstract sense, totally or partially superposed with each other and simultaneously (depending on the purposes of the investigation) be provided, to one degree or another, with the ability 'to interact' with other bodies (for example, with original bodies being the base for the construction). Due to the above-said, the set $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$ should be understood as a potentially (not actually) existing set. This fact adds flexibility to the GIRRM proposed by Rogovtsov and Samson (1976); Rogovtsov (1980, 1981a, 1981b, 1981c, 1983, 1985a, 1985b, 1985d, 1989, 1999).

One of the most important steps that should be made while designing this general method was the construction of a set of unary algebraic operations that are defined on the set $\{V_\beta\}_{\beta\in B}$ and leave radiation fields in turbid bodies invariant almost everywhere (these bodies are the elements of set $\{V_\beta\}_{\beta\in B}$). Here, the unary algebraic operations stands for the mappings of set $\{V_\beta\}_{\beta\in B}$ into the set $\{V_\beta\}_{\beta\in B}$ itself (the definition of the unary algebraic operations is given, for example, in Faure et al., 1964; also see the Appendix). There were introduced the under-mentioned sets of unary algebraic operations (each of which should be made instantaneously) on set $\{V_\beta\}_{\beta\in B}$ by Rogovtsov (1980, 1981a, 1981b, 1989, 1999).

(i) Section of each body $V_\beta \in C$ (C is any non-empty subset of the set $\{V_\beta\}_{\beta\in B}$) by surface $(\Xi_j \cap V_\beta)$, which does not introduce additional discontinuities in physical characteristics to the discontinuities already existing in V_β. The elements belonging to $\{V_\beta\}_{\beta\in B}\setminus C$ remain unchanged.

(ii) Replacement of the effect of boundary S_β of any body V_β ($V_\beta \in D$, where D is any non-empty subset of the set $\{V_\beta\}_{\beta\in B}$) on radiation field in V_β by the effect of other boundary, which coincides with S_β in the geometrical sense and on which absorbers and radiators are located. The absorbers totally absorb radiation incident on them from any direction. The radiators emit the same radiation as has been implemented in original body V_β and as propagates from $S_\beta(V^0_\beta)$ at any point, selected on this surface, in the directions specified by the rays, emerging from the surface and intersecting all the regions $V^0_\beta \cap O$ (O is any neighbor of the given point in E^3). With this, the elements belonging to $\{V_\beta\}_{\beta\in V}\setminus D$ remain unchanged.

(iii) Replacement of the effect of portion $S^\times_\beta\left(S^\times_\beta \subset S_\beta\right)$ of boundary S_β of any body V_β ($V_\beta \in H$, where H is any non-empty subset of the set $\{V_\beta\}_{\beta\in B}$) on radiation field in V_β by the effect of other surface S^*_β, which coincides with $S^\times_\beta$ in the geometrical sense. At this, there are absorbers and radiators on surface S^*_β. The absorbers totally absorb radiation incident on them from any direction. The radiators emit the same radiation as has been implemented in original body V_β and as propagates from $S_\beta(V^0_\beta)$ at any point, selected on this surface, in the directions specified by the rays, emerging from the surface and intersecting all the regions $V^0_\beta \cap O$ or $(E^3\setminus V_\beta)\cap$ O (O is any neighbor of the given point in E^3). The elements of set $\{V_\beta\}_{\beta\in B}\setminus H$ remain unchanged.

(iv) Any ordered sequence of the operations defined by items (i) to (iii).

Let the above statements be additionally simplified. The standard notations applied in mathematics (see, for example, Fraenkel and Bar-Hillel, 1958; Faure et al., 1964; Kuratowski and Mostowski, 1967; also see the Appendix) were used, while stating the unary algebraic operations of items (i) to (iv). In particular, expression $A \cap B$ has the sense of the operation result of the intersection of sets A and B, and symbol $A \subset B$ has the sense of the inclusion relation of set A into set B (see additional explanations in the Appendix). Assume that, after the section of body V_β by surface $(V_\beta \cap \Xi_j)$, the points of the latter do no belong to the body obtained as a result of operation (i). At this, a new body can generally differ by its topological properties from V_β. It should be noted that the equal unary algebraic operations of types (i) to (iv) are stood for operations that map equal

bodies to equal ones. Absorbers and radiators should be realized as mental objects located at places of the replaceable boundaries and 'infinitesimally distanced' from the sides of simple pieces of the surfaces, from which these boundaries (or their portions) are composed. Conceptions of *absorbers* and *radiators* introduced above have only an abstract sense. This is due to their constructive implementation as the materialized objects is impossible. To understand this, it is sufficient to recall that a physical body absorbing incident radiation totally will necessarily emit radiation, i.e. it cannot generally be an absorber in the above sense. However, mental operations and experiments carried out using these conceptions lead to valid and mathematically correct results enabling one to get nontrivial consequences. More detailed explanations and evidences in this respect were given by Rogovtsov (1980, 1981a, 1981b, 1981c, 1983, 1985d, 1989, 1999). These publications also provided the detailed explanation of the ways one could constructively utilize unary algebraic operations (i)–(iv). In particular, it was pointed out that, using these operations, one could geometrically and physically isolate parts of a turbid medium from an original turbid medium without destroying the characteristics of radiation fields in regions outside original and artificially made boundaries. Besides, there were constructed the general procedures connecting operations of physical isolation and redefinition in the said works. Just these procedures enable one to correctly make the imbedding of turbid media of complex configurations in other, more volumetric turbid media having simpler shapes.

Fig. 7.1 illustrates one of the simplest examples of the action of operation (i) on set $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$, when one takes a single its element (i.e. a single body V_β; $C = \{\mathrm{V}_\beta\}$) as a subset C. Solid arrows and lines limiting the arrows show schematically angular patterns of the radiation intensity that goes through some arbitrarily selected points (they coincide with the centers of the black circles) lying on boundary S_β (or $\mathrm{S}_{\beta'}$) and inside body V_β (or $\mathrm{V}_{\beta'}$), respectively. Symbol $(\Xi_j \cap \mathrm{V}_\beta)$ is meant as a surface that is the intersection of some surface Ξ_j with body V_β. The left and right parts of Fig. 7.1 coincide with each other in physical and geometrical senses, except that there is an additional surface $(\Xi_j \cap \mathrm{V}_\beta)$ in the right part (this means that boundary $\mathrm{S}_{\beta'}$ of body $\mathrm{V}_{\beta'}$ is a union of boundary S_β and surface $(\Xi_j \cap \mathrm{V}_\beta)$, i.e. $\mathrm{S}_{\beta'} = \mathrm{S}_\beta \cup (\Xi_j \cap \mathrm{V}_\beta)$. It should be noted that the two parts of Fig. 7.1 are artificially separated to better understand the essence of operation (i). In reality, bodies V_β and $\mathrm{V}_{\beta'}$ occupy the same space portion with the accuracy of the differences between boundaries S_β and $\mathrm{S}_{\beta'}$. Note once more that the centers of the black circles in Fig. 7.1 can be at any points lying on S_β and $(\Xi_j \cap \mathrm{V}_\beta)$, respectively.

Fig. 7.2 illustrates the action of operation (ii) on set $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$, when one takes only a single body V_β (i.e. $D = \{\mathrm{V}_\beta\}$) as a subset D. The bold line in the right-hand part of Fig. 7.2 represents absorbers and boundary (in geometrical sense) of body V_{β_1}. The dotted line in the right-hand part of Fig. 7.2 shows schematically *radiators* (they are on the inner portion of boundary S_{β_1}, i.e. on $\mathrm{S}_{\beta_1}\big(\mathrm{V}^0_{\beta_1}\big)$). For clearness, the dotted line is noticeably moved to the interior of body V_{β_1} (in reality, this line should be at the infinitesimal distance from the bold line, i.e. from the absorbers). It should be especially noted that, in physical and geometrical senses, the interior portions of bodies V_β and V_{β_1} coincide (i.e. $\mathrm{V}^0_\beta = \mathrm{V}^0_{\beta_1}$), although they are separated from each other in Fig. 7.2. Solid arrows and lines limiting the arrows in the both parts of Fig. 7.2 have the same sense as in Fig. 7.1.

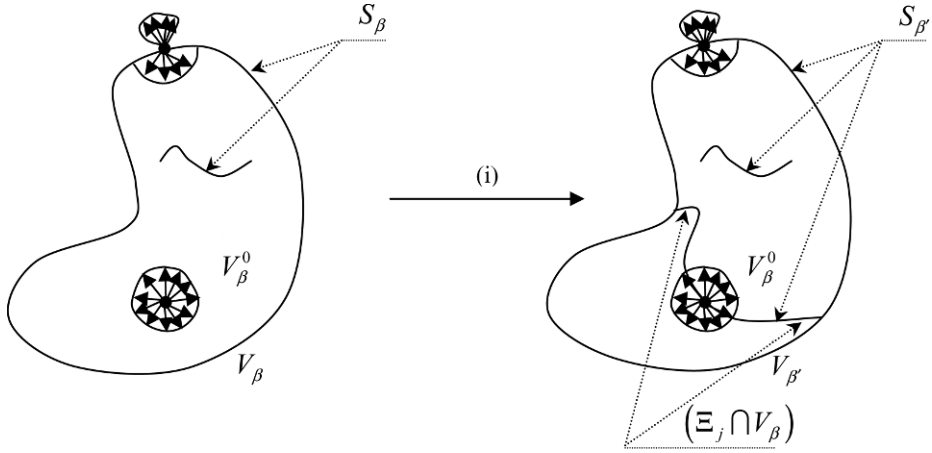


Fig. 7.1. Example of operation (i)

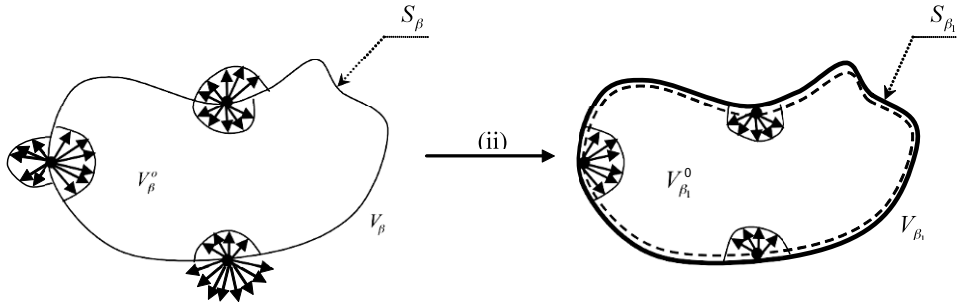


Fig. 7.2. Example of operation (ii)

Fig. 7.3 illustrates one of the simplest examples of the action of operation (iii) on the $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$, when one takes a single its element (i.e. a single body V_β; $H = \{\mathrm{V}_\beta\}$) as a subset H. Solid arrows and lines limiting the arrows in both parts of Fig. 7.3 have the same sense as in Fig. 7.1 and Fig. 7.2. The bold line in the right-hand part of Fig. 7.3 represents absorbers. The dotted line in the right-hand part of Fig. 7.3 shows schematically radiators. For clearness, the dotted line is noticeably moved to the interior of body $\mathrm{V}_{\beta''}$ (in reality, this line should be at the infinitesimal distance from the bold line, i.e. from S^*_β). In physical geometrical, the interior portions of bodies V_β and $\mathrm{V}_{\beta''}$ coincide (i.e. $\mathrm{V}^0_\beta = \mathrm{V}^0_{\beta''}$), although they are separated from each other in Fig. 7.3.

Let a set of the above-described unary algebraic operations on $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$ be denoted by $\mathfrak{R}$. Elements $r \in \mathfrak{R}$ specify the following mappings of r: $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}} \to \{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$. Set $\mathfrak{R}$ is a semi-group with respect to a binary algebraic operation of the composition of mappings. Operation (iii), when an empty set is taken as portion $\mathrm{S}^\times_\beta$ boundary S_β of any body $\mathrm{V}_\beta \in \{\mathrm{V}_\beta\}_{\beta\in B}$, is an unit (neutral) element e that coincides with the identical mapping (that does not change the elements of set $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$). The general analysis of mathematical properties of the solutions to

boundary-value problems of the radiative transfer theory as well as physical considerations and conceptions used in this theory unambiguously confirm the validity of the following general invariance principle (Rogovtsov (1981a, 1981b, 1989, 1999)):

The radiation field in any bounded subset of each $\mathrm{V}_\beta \in \{\mathrm{V}_\beta\}_{\beta\in B}$ *is almost everywhere invariant with respect to unary algebraic operations being the elements of semi-group* $\mathfrak{R}$.

It should be especially underlined that in the overwhelming majority of scientific publications, where the classic conceptions on invariance (covariance) and on symmetry were utilized for solving particular or general problems of physics, mathematics, technology, etc., the invariance and symmetry principles stood for general statements establishing the invariance (i.e. the constancy in some specific or abstract sense) of certain objects, laws, systems, constructions, art and architecture products, etc. with respect to sets of algebraic operations forming a group (in particular, the Lie group). The fundamental role, which was played by the classic conceptions on invariants, invariance and symmetry principles in mathematics, physics, and art as well as specific applications of the conceptions for solving various problems, was thoroughly analyzed, for example, by Klein (1926); Weyl (1952, 1968); Einstein (1965, 1966); Wigner (1970); Ovsyannikov (1982); Ibragimov (1985); Bogolyubov and Shirkov (1976); Rogovtsov (1999); Kovalev and Shirkov (2008).

The general invariance principle stated above does not allow a treatment similar to a classic invariance principle establishing the invariance of some objects, quantities, etc. with respect to a set of operations that is a group (Lie group). This is due to the fact that semi-group $\mathfrak{R}$ cannot be imbedded into a group (the evidence for this statement was given by Rogovtsov (1981b, 1989, 1999)). It should be noted that the GIP refers to the statements, in which the invariance of physical and mathematical objects, systems, constructions, etc. with respect to operations having the sense of 'partitions', 'redefinitions', and 'imbeddings' (the classic invariance principles sometimes partially use operations of the 'partition' kind too). It

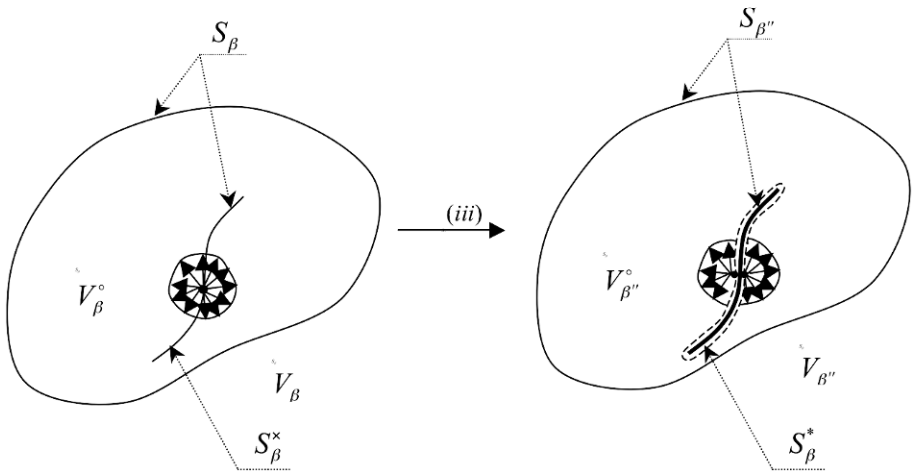


Fig. 7.3. Example of operation (iii)

follows from the above-said that the properties and the invariance principles should be generally understood as the statements on the invariance or partial invariance of some set M (of objects, systems, constructions, solutions, etc.) with respect to some set W of operations (actions) that can be provided with certain algebraic properties (however, W needs not to necessarily be a group or even a semi-group). More detailed explanations of this topic were made by Rogovtsov (1999).

The GIP, as any formalized and balanced system of postulated statements too, comprises a considerable part of the information (partially in implicit form) that should be gathered by making various constructions and conclusions. At this, it is rather useful to utilize the physical conceptions of ray optics and the probabilistic treatment of the radiative transfer theory (Sobolev, 1951; Ueno, 1957, 1958; Minin, 1988). Let a consequence of the GIP, which can be derived by these physical conceptions and probabilistic treatment, be stated. This consequence is essentially a property of the invariance. Consider some body $\mathrm{V}_\beta \in \{\mathrm{V}_\beta\}_{\beta\in B}$. At a certain time moment $t = t_0 - 0$, apply operation (ii) to the body V_β, when $\mathrm{D} = \{\mathrm{V}_\beta\}$ (since D is a set, then symbol V_β is posed in the curly brackets according to the notation used in the theory of sets). Due to the operation, V_β will be mapped into another body $\mathrm{V}_{\beta_1} \in \{\mathrm{V}_\beta\}_{\beta\in B}$, which differs from V_β only in the presence of absorbers and radiators, but there is a valid equality $\mathrm{V}^0_\beta = \mathrm{V}^0_{\beta_1}$ inside the body V_β (the equality is understood in the geometrical and physical senses). Assume mentally that the absorbers are instantaneously removed from V_{β_1} at the time moment $t = t_0 - 0$ without destroying the structure of the turbid medium and to radiation field (the boundary of the body becomes completely transparent for radiation after this). Thereafter, the body V_{β_1} will be mapped onto the body V_{β_2}, the boundary of which transmits all the incident radiation without any changes, and $\mathrm{V}^0_{\beta_2} = \mathrm{V}^0_\beta$ from the viewpoints of optical properties of the medium, presence of internal sources in $\mathrm{V}^0_{\beta_2}$, and initial (at time moment $t_0 - 0$) distribution of the radiation field. However, the optical properties of the boundaries of the bodies V_β and V_{β_2} can differ from each other. Besides, there are the radiators on $\mathrm{S}_{\beta_2}\big(\mathrm{V}^0_{\beta_2}\big)$. At $t > t_0$, the radiation field in $\mathrm{V}^0_{\beta_2}$ can change, if the photons intersecting S_{β_2} are able to again return back to $\mathrm{V}^0_{\beta_2}$. Consider now at some time moment t, at any point of $\mathrm{V}^0_{\beta_2}$, the whole set of photons that were emitted, during the time duration from $t_0 - 0$ to t, by the radiators and by the internal sources of $\mathrm{V}^0_{\beta_2}$ or that were in flight at $t_0 - 0$ and did not intersect the boundary S_{β_2} during their random walk. The non-intersection of a boundary of some body will be everywhere below understood as the absence of the reflection from the boundary or of the transmission by it. The removal of other photons from consideration is equivalent to the operation that is made by an absorber, which removes all the photons incident on it. The validity of the statement follows from the above discussion and from the GIP (Rogovtsov, 1981a, 1999):

That part of the radiation field, which is being formed or has been formed due to the random walk of photons emitted, during the time interval from $t_0 - 0$ *to* t*, where* $t > t_0$*, by the radiators and the internal sources or being in flight in* V^0_β *at time moment* $t = t_0 - 0$ *exclusively inside the interior* V^0_β *of body* V_β*, coincides with the true radiation field implemented in* V^0_β*.*

This property of the invariance will be utilized in subsection 7.2.3 for deriving the general invariance relations (GIRs).

Now the scientific publications, in which important results have been obtained to be, to a certain degree, the original basis for stating the GIP and constructing the GIRRM, will be pointed out. Also, the papers, in which the particular versions of the GIP have been stated or announced and the formal interrelations between the problems of the radiative transfer theory and other topics of physics and mathematical physics have been noted, will be indicated. The first fundamental investigations of the RTT problems for the case of plane-parallel and exclusively one-dimensional media have been fulfilled by Ambartsumian (1943a, 1943b, 1944a, 1944b, 1947, 1948) using explicitly the invariant properties of radiation field characteristics with respect to certain operations. The first invariance principle for the RTT has been formulated as two statements in the papers (Ambartsumian, 1943a; 1943b) to get, on this basis, the nonlinear integral equations for the reflection function (for a semi-infinite plane-parallel macroscopically homogeneous medium and arbitrary phase function) and the auxiliary functions, through which the reflection function is expressed. Moreover, the analytical relations for the reflection and transmission functions of a plane-parallel slab with a finite optical thickness have been written down in the paper (Ambartsumian, 1943a) for the case of isotropic scattering and the nonlinear integral equations for corresponding auxiliary functions have been stated. Nontrivial applications of the said invariance principle to the solutions to some boundary-value problems of the RTT for the cases of a plane-parallel medium with large optical thickness and of a semi-infinite plane-parallel medium have been given by (Ambartsumian (1944a, 1948). However, the papers (Ambartsumian, 1943a, 1943b, 1944a, 1948) used explicitly only the invariance properties of radiation field characteristics at the boundaries of macroscopically homogeneous plane-parallel media with respect to the operations of addition (or addition and subtraction) of layers with a finite or an infinitesimal optical thickness. It should be especially noted that the same papers by Ambartsumian (1944b, 1947) utilized, besides the above-said invariance properties, essentially (but generally implicitly) the particular versions of the operations of kinds (i) and (ii) defined in subsection 7.2.2. The constructive procedures exploited in all the papers by Ambartsumian (1943a, 1943b, 1944a, 1944b, 1947, 1948) were then combined under the general name of the 'layer adding method'. The essence of this method was explained in detail by Ambartsumian (1944b and 1947) using rather simple and instructive examples (they refer to the RTT problems for the case of exclusively one-dimensional isotropic and non-isotropic media). Really, Ambartsumian's papers first showed, how, on the basis of using the invariance properties (in explicit or implicit forms), one can state and solve functional equations (in particular, integral ones) for desired radiation field characteristics at the boundaries of exclusively one-dimensional or plane-parallel media with any optical thickness. It should be noted, as the merit of the approach proposed and implemented by Ambartsumian while solving some boundary-value problems of the RTT for the case of macroscopically homogeneous plane-parallel media, that this approach provided its evolution and generalization over a wide variety of directions. One can especially isolate the paper (Stokes, 1862) among the earlier publications, where the constructions similar to the layer adding method have been used. The problem of deriving the reflection and transmission coefficients of a pile of the same plates (here an absorbing homogeneous medium is supposed to be between the plates) by the known respective coefficients of a single

plate has been solved there. In spite of the striking resemblance of the functional equations for the reflection and transmission coefficients obtained by Stokes (1862) and by Ambartsumian (1944b, 1947) while solving different problems, the considerations used by them have differed from each other essentially. In the paper by Stokes (1862), the functional equations were derived by means of successive treatment and summation of the contributions of all the events of light interaction (their number was infinite) with each of the two subsystems of the plate piles, by which the original plate pile were partitioned, to the reflection and transmission coefficients. In the papers of Ambartsumian (1944b, 1947), the invariance properties of the radiation field, having formed in an exclusively one-dimensional medium, with respect to the particular cases of the operations of kinds (i) and (ii) (see subsection 7.2.2) were partially used (although in the implicit form) to derive the functional equations. Besides, in the said works by Ambartsumian, there has been constructively used the information on the radiation field characteristics in a composite media formed by combining two one-dimensional media. It should be noted that some starting ideas being the foundation of the classical version of the invariant imbedding method have been essentially outlined in a particular form in the papers by Stokes (1862) and Ambartsumian (1944b, 1947). The substantial contribution to the development of the basic ideas and constructions of this version of the invariant imbedding method was made by Chandrasekhar (1950), Redheffer (1954), Bellman (1954), Bellman and Kalaba (1956) and Sobolev (1956). First nontrivial applications of this method were given in the papers by Bellman and Kalaba (1956) and Sobolev (1956), who considered the radiative transfer process through macroscopically inhomogeneous plane-parallel media. The paper (Bellman and Kalaba, 1956) has stated, in rather general form, the classical version of the invariant imbedding principle on the base of more abstract treatment of the ideas and the constructions used in the above-mentioned publications. One can find various applications and reviews of the papers, where the principle has been utilized, in the publications (for example, Angel and Bellman (1972), Bellman (1986), Klyatskin (1986) and Rogovtsov (2008)). After the investigations performed by Ambartsumian on the basis of using the invariance principles and the layer adding method stated by him, the next step was made by Chandrasekhar (1950). Remaining within the frame of the consideration of RTT problems, Chandrasekhar has stated four invariance principles. These principles establish essentially the validity of four balance relations, which connect the radiation intensity (generally, at optical thickness τ inside a plane-parallel slab with optical thickness τ_0) for the case of external mono-directional illumination with the reflection or transmission coefficients, which can correspond to plane-parallel slabs with, generally, different optical thickness (this thickness can assume one value from the set of τ, τ_1, $\tau_1 - \tau$). The relations derived by Chandrasekhar (1950) can be treated as the invariance relations being the consequences of not only the definition of the reflection and transmission functions (Chandrasekhar, 1950), but also of the invariance of the radiation intensity in a plane-parallel medium with respect to the particular operations of kinds (i) and (ii) (see subsection 7.2.2). It should be noted that, in the monograph (see Chandrasekhar, 1950), there have been first obtained the invariance relations comprising the radiation field characteristics inside macroscopically homogeneous plane-parallel media. Moreover, Chandrasekhar first showed the applicability of the usage of the invariance relations for solving

boundary-value problems for the vector radiative transfer equation. It should be added to the above-said that he first derived the equation set for the reflection and transmission functions comprising first partial derivatives from these functions with respect to the optical thickness of a slab (i.e. he was the first who obtained the first important constructive result referring essentially to the statement of the starting equations of the invariant embedding method). The paper (Kadomtsev, 1957) made the first attempt to reformulate the invariance principle of Ambartsumian (1943a) for the case of a macroscopically homogeneous medium of an arbitrary geometrical shape. After the classical publications by Ambartsumian and Chandrasekhar, many researchers on the basis of these relations over about three decades have not only obtained more general invariance relations (they are often called invariance principles), but also proposed effective analytical and numerical methods for calculating the reflection and transmission functions and the characteristics of internal radiation fields in plane-parallel media (not necessarily macroscopically homogeneous), which can comprise internal radiation sources. Among these publications, see in particular (van de Hulst, 1963, 1980; Sobolev, 1968, 1975; Shimizu, 1968; Hansen, 1969; Hunt and Grant, 1969a, 1969b; Yengibarian and Mnatsakanian, 1974; Ivanov, 1975; Ivanov and Volkov, 1978; Domke, 1976, 1978a, 1978b; Yanovitskij, 1979, Rogovtsov, 1980; Domke and Yanovitskij, 1981, Viik, 1982; Nikoghossian and Haruthyunian, 1989; Yengibarian and Melkonian, 1989). More detailed bibliographical references can be found, for example, in the monographs (Yanovitskij, 1997; Rogovtsov, 1999). All the above-mentioned publications used explicitly or implicitly the invariance of radiation fields with respect to particular cases of the operations of (i) and (ii) (see subsection 7.2.2). Moreover, the papers (Ivanov, 1975; Rogovtsov, 1980; Yanovitskij, 1979, 1981) utilized once more the property of the invariance (it is the consequence of the GIP). The essence of this invariance property is substantially reduced to the statement on the invariance of the radiation field in an isolated sublayer of an original plane-parallel slab with respect to the particular version of the operation of kind (ii) and on the equality of the radiation field in the isolated part (sublayer) with the corresponding external sources ('radiators') at its boundaries to the radiation field at respective points of the original slab. If one looks more abstractly at the considerations and constructions utilized in the above-mentioned works, it can be concluded that therein have been essentially used the operations of the isolation of a part of an object from a whole object and the operations of embedding and redefinition. These operations in the substantially more general form are the foundation of the GIP. Note two additional important aspects that differ from the approach based on the usage of the invariance properties in the radiative transfer theory from other methods. One of these aspects was mentioned by Bellman and Kalaba (1956); Yengibarian and Mnatsakanian (1974). Its sense is that, even remaining within the frame of the consideration of RTT problems for the case of plane-parallel media, one can establish the presence of the relationships of this approach with the semi-group theory. The second aspect was discussed in detail by Mnatsakanian (1982), Shirkov (1989), Rogovtsov (1999) and Kovalev and Shirkov (2008). It is related with the formal resemblance, established and evaluated in the papers (Mnatsakanian, 1982; Shirkov, 1989), between the equations derived in the nonlinear radiative transfer theory by means of the above-said approach and

the equations, which are used in the theory of interacting quantized fields within the frame of the renormgroup approach (Kovalev and Shirkov, 2008).

In a number of works (Kadomtsev (1957), Rogovtsov (1989), Nikoghossian (2009)) there were pointed out and stated connections among Ambartsumian's invariance principle and variation principles and a group operation (translation). These connections exist for the case of macroscopically homogeneous media. In Nikoghossian (2009) it is asserted that Ambartsumian's invariance principle is a consequence of a variation principle whose formulation uses Lagrangian density not depending explicitly on the coordinates of an observation point. In addition Ambartsumian's invariance principle is considered by Nikoghossian as a conservation law. This law has been obtained using a generalization of Nöether's theorem for the case of integro-differential equations. It should be noted that in Ambartsumian (1943a, 1943b, 1944a, 1944b, 1947, 1948), in solving particular boundary-value problems of RTT for the case of plane parallel media on the basis of Ambartsumian's invariance principle and the layer adding method, a more essential role was played by the use of particular cases of semi-group operations of the sort (i) and (ii), although in a implicit form. It was this general special feature of Ambartsumian's invariance principle and the layer adding method that was one of points in formulating GIP and the derivation of GIRs. It should be noted once more that constructing GIRRM on the basis of GIP and GIRs allows one to solve boundary-value problems of RTT for the case of homogeneous and inhomogeneous dispersion media of different configurations. These media can be limited to underlying surfaces and can contain underlying surfaces in it. In addition their configuration and optical characteristics cannot possess any substantial symmetry. It is important that the GIRRM method allows one to obtain different integral relations (conversation laws) for the case of absence of the symmetry pointed out. Examples of such integral relations (some of them are vector integral relations) were obtained in works (Rogovtsov, 1993, 1994, 1999). However in the derivation of these integral relations in some cases it is necessary to use the homogeneous property of the point Euclid space (in fact it is a symmetry property) and the independence of their structure on the selection of a coordinate system. In other words, although the use of semi-group operations is a basis of GIP and GIRRM, it is very important in solving boundary-value problems of RTT to take into account internal symmetries in stating these problems and symmetry properties of coordinate systems used.

7.2.3 General invariance relations and their physical interpretation

The scheme of the derivation and of the application of the general invariance relations (GIRs) for solving boundary-value problems of the radiative transfer theory (RTT) will now be described. These relations are the consequences of the GIP. The sequence of the considerations that will be utilized for deriving the GIRs is essentially independent (except for the outward details) of the physical essence of single absorption, scattering and emission events. It is, however, required that the conditions for the applicability of ray optics conceptions and of the probabilistic interpretation of the radiative transfer process (see, for example, Rozenberg, 1977; Sobolev, 1951; Ueno, 1957, 1958) will be satisfied. The scheme of the derivation of the GIRs proposed by Rogovtsov (1980, 1981a, 1981b, 1981c, 1983, 1985a, 1985b, 1985c, 1989,

1999) includes a procedure of the papers (Pikichian, 1982a, 1982b, 1983, 1989) as a particular case. Procedures enabling one to get consequences from the GIRs themselves should be also utilized, while solving specific boundary-value problems of the radiative transfer theory. These procedures were constructively used by Rogovtsov (1980, 1981a, 1981c, 1983, 1985a, 1985b, 1985c, 1986a, 1986b, 1986c, 1988a, 1988b, 1988c, 1988d, 1990a, 1990b, 1991a, 1991b, 1991c, 1992a, 1992b, 1993, 1997, 1999, 2002, 2007); Pikichian (1982b, 1983, 1989); Rogovtsov and Samson (1985a, 1985b, 1987); Anisimov and Rogovtsov (2002); Rogovtsov and Karpuk (2003), while solving analytically or semi-analytically specific, rather complex problems of the RTT for turbid media of various configurations. Besides, more abstract treatment of the said schemes and procedures (in the spirit of the general approach proposed by Rogovtsov, 1999) has enabled a number of complex problems of mathematical physics to be solved in an analytical form. In particular, the Cauchy problem for an infinite set of linear differential equations, which describes the collision-free process of the molecule excitation (Rogovtsov, 1990c), has been solved in a analytical form, and the solution to a family of nonlinear boundary-value problems for ordinary differential and integro-differential equations has been reduced to the solution to the Cauchy problem (Rogovtsov (2008)). Moreover, the usage of the ideas of the approach elaborated by Rogovtsov (1980, 1981a, 1981b, 1981c, 1983, 1985a, 1985b, 1985d, 1999) enabled one to get the solution to the characteristic equation of the RTT (Rogovtsov, 1996; Rogovtsov and Borovik, 2009) in the analytical form for a case of an arbitrary phase function. Due to the above, the said scheme and procedures can be regarded as rather substantial and constructive ones.

The scheme of the derivation of the GIRs is illustrated by an example of monochromatic scattering within the scope of the scalar radiative transfer theory. Let all the above-mentioned assumptions on the geometrical properties of the boundaries of turbid bodies, the optical and geometrical properties of the interiors of these bodies and of underlying surfaces located therein be additionally assumed to be fulfilled. Note that the scheme outlined below of the derivation of the GIRs is formally applicable too for cases where the polarization characteristics of radiation and the frequency redistribution processes should be taken into account.

7.2.3.1 Derivation of the general invariance relations for the case of non-stationary sources

Let there be radiation sources inside body $\mathrm{V}_\beta \in \{\mathrm{V}_\beta\}_{\beta\in B}$, the density of which is specified by function $g_\beta\big(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta^0\big)$, and the radiation intensity in V_β^0 be $I\big(\vec{r},\vec{\Omega},t_0-0;\mathrm{V}_\beta\big)$ at time moment $t=t_0-0$ ($t_0\in(-\infty,+\infty)$). Assume that local optical characteristics of turbid medium V_β^0 are independent of the radiation field, but can be dependent on the radiation field on boundary S_β and outside V_β. At time moment $t=t_0-0$, apply operation (ii) to $\{\mathrm{V}_\beta\}_{\beta\in B}$ with $\mathrm{D}=\{\mathrm{V}_\beta\}$. The radiation field does not change with this (see the GIP), but body V_β will be mapped to object $\mathrm{V}_{\beta_1}(\beta_1\in\mathrm{B})$ that differs from V_β only in the boundary properties. Let body V_{β_2} be V_{β_1}, from which the absorbers are removed at time moment $t=t_0-0$, and there is 'positioned' an underlying or another surface S_{β_2} with specified linear local reflection and transmission operators instead of the absorbers. From $\{\mathrm{V}_\beta\}_{\beta\in B}$ one can always select a body V_{β_3} that contains a part identical to

V_{β_2} from the viewpoint of the properties of and the sources of the medium. Assume for simplicity that $g_{\beta_3}\left(\vec{r},\vec{\Omega},t;\mathrm{V}^0_{\beta_3}\right)\equiv 0$ outside of the part V_{β_2} of body V_{β_3}. At this, in part $\mathrm{V}_{\beta_3}\backslash\mathrm{V}_{\beta_2}$, the properties of body V_{β_3} are independent of the radiation field. Moreover, let there be no radiation field in part $\mathrm{V}_{\beta_3}\backslash\mathrm{V}_{\beta_2}$ of body V_{β_3} at time moment $t=t_0-0$ and the body V_{β_3} itself be not irradiated by external radiation at time interval from t_0-0 to $+\infty$. Apply then operation (i) with $\mathrm{C}=\{\mathrm{V}_{\beta_3}\}$ and $\Xi_j=\mathrm{S}_{\beta_2}$ to $\mathrm{V}_\beta\in\{\mathrm{V}_\beta\}_{\beta\in B}$ (the equality is understood in the geometrical sense only). This operation partitions V_{β_3} by three subsystems $\mathrm{V}^0_{\beta_2}$, S_{β_2}, and $\mathrm{V}_{\beta_3}\backslash\mathrm{V}_{\beta_2}$.

Radiation intensity $I\left(\vec{r},\vec{\Omega},t;\mathrm{V}_{\beta_3}\right)$ in body V_{β_3} can be represented in two ways, if one partitions the photon 'trajectories' by different topological types and accounts for the 'interaction' of these subsystems. Note that the presence of an initial radiation field in part V_{β_2} of body V_{β_3} at time moment $t=t_0-0$ is equivalent at $t>t_0$ to specifying the density of internal radiation sources of kind $\Theta_{\mathrm{V}^0_\beta}(\vec{r})\,I\left(\vec{r},\vec{\Omega},t_0-0;\mathrm{V}_\beta\right)\left(v(\vec{r})\right)^{-1}\delta(t-t_0)$, where $\Theta_{\mathrm{M}}(\vec{r})$ is the characteristic function of set M ($\Theta_{\mathrm{M}}(\vec{r})$ is unity, when $\vec{r}$ specifies points of M, and is zero otherwise). Radiation intensity $I\left(\vec{r},\vec{\Omega},t;\mathrm{V}_{\beta_3}\right)$ in body V_{β_3} at $t>t_0$ is completely defined by the internal radiation sources in $\mathrm{V}^0_{\beta_2}$, which are described by density $g^{\times}\left(\vec{r},\vec{\Omega},t;\mathrm{V}_3\right)=\Theta_{\mathrm{V}^0_\beta}(\vec{r})\left[g_\beta\left(\vec{r},\vec{\Omega},t;\mathrm{V}^o_\beta\right)+I\left(\vec{r},\vec{\Omega},t_0-0;\mathrm{V}_\beta\right)\left(v(\vec{r})\right)^{-1}\delta(t-t_0)\right]$, by the radiators located on $\mathrm{S}_{\beta_2}\left(\mathrm{V}^0_{\beta_2}\right)$, and by volume Green function $G_*\left(\vec{r},\vec{\Omega};\vec{r}',\vec{\Omega}';t;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right)$ of the radiative transfer equation of Eq. (7.1) for the case of body V_{β_4}. Body V_{β_4} stands for body V_{β_3}, from which the radiators and the internal sources determined by density $g^{\times}(\ldots)$ are removed at time moment $t=t_0-0$. Symbol S_{β_2} is introduced explicitly in the notation for the volume Green function $G_*(\ldots)$, since there will be considered various situations, in which the properties of S_{β_2} should be changed. The intensity of the radiation field in V_{β_3} can be also represented, at $t>t_0$, as a sum of terms $\left(\mathrm{I}_1\left(\vec{r},\vec{\Omega},t;\mathrm{V}_{\beta_3}\right)+\mathrm{I}_2\left(\vec{r},\vec{\Omega},t;\mathrm{V}_{\beta_3}\right)\right)$. To do so, the 'interaction' of the above-said subsystems should be considered. The first term $\mathrm{I}_1\left(\vec{r},\vec{\Omega},t;\mathrm{V}_{\beta_3}\right)$ is equal to $\Theta_{\mathrm{V}^0_\beta}(\vec{r})\,I\left(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta\right)$. It coincides with the portion of the radiation intensity in V_{β_3} that is due to photons emitted by the radiators and the internal sources in body V_{β_2} (their density is specified by function $g^{\times}\left(\vec{r},\vec{\Omega},t;\mathrm{V}_3\right)$) and not intersecting the boundary S_{β_2} during the random walk until they come to an 'observation' point in $\mathrm{V}^0_{\beta_2}$ (it follows from the GIP and the invariance property stated in subsection 7.2.2). The second term $\mathrm{I}_2\left(\vec{r},\vec{\Omega},t;\mathrm{V}_{\beta_3}\right)$ gives the portion of the contribution to the total radiation field from photons, which are formed at $t>t_0$ in V_{β_3} due to the exit outside of $\mathrm{V}^0_{\beta_2}$ and subsequent random walk in the whole body V_{β_3} and which initially generate the term $\mathrm{I}_1\left(\vec{r},\vec{\Omega},t;\mathrm{V}_{\beta_3}\right)$. By equalizing the expressions for the radiation intensity $\mathrm{I}\left(\vec{r},\vec{\Omega},t;\mathrm{V}_{\beta_3}\right)$ obtained in the above two ways, one concludes the following general invariance relation (GIR) (Rogovtsov, 1981a, 1981b, 1989, 1999):

$$
\begin{aligned}
&\Theta_{\mathrm{V}_\beta^0}(\vec{r})\, I\left(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta\right)\\
&=\iiint\limits_{\mathrm{V}_\beta^0} d\mathrm{V}' \int\limits_{\Omega} d\Omega' \int\limits_{t_0-0}^{t} G_*\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}';t-t';\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right) g_\beta\left(\vec{r}',\vec{\Omega}',t';\mathrm{V}_\beta^0\right) dt'\\
&-\iint\limits_{\mathrm{S}_\beta(\mathrm{V}_\beta^0)} dS' \int\limits_{\Omega} \left(\vec{n}'\cdot\vec{\Omega}'\right) d\Omega' \int\limits_{t_0-0}^{t} G_*\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}',t-t';\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right) I\left(\vec{r}',\vec{\Omega}',t';\mathrm{V}_\beta\right) dt'\\
&+\iiint\limits_{\mathrm{V}_\beta^0} d\mathrm{V}' \int\limits_{\Omega} G_*\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}';t-t_0;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right) \left(v\left(\vec{r}'\right)\right)^{-1} I\left(\vec{r}',\vec{\Omega}',t_0-0;\mathrm{V}_\beta\right) d\Omega',\\
&t>t_0. \qquad (7.16)
\end{aligned}
$$

This GIR is valid for any $\vec{\Omega}\in\Omega$ and any $\vec{r}$ specifying points in body V_{β_4} not belonging to S_β and to the underlying surfaces (at these surfaces, the characteristics of the radiation field can have no clear physical sense).

Equation (7.16) is one of the most common GIRs in the integral form (from a viewpoint of structure, not of accounting for the comprehensive laws of radiation interaction with the medium). For a plane-parallel turbid medium, the considerations similar to the above, while deriving the GIR of Eq. (7.16), were first utilized by Rogovtsov (1980). Another GIR that is one of the most common ones was derived by Rogovtsov (1981b, 1999) using unary operation (iii) (see the GIP). This GIR enables one to reduce the original boundary-value problem for the SRTE of Eq. (7.1) to a more simple boundary-value problem, if one knows the intensity at all the sides of the simple pieces composing portion $\mathrm{S}_\beta^\times$ of boundary S_β of body V_β. Notice particularly that the said GIRs connect the solutions to different or similar boundary-value problems for the SRTE of Eq. (7.1). Since one can take any elements from set $\{\mathrm{V}_\beta\}_{\beta\in B}$ and specify the properties of S_β, S_{β_2}, V_β, V_{β_4}, etc. as necessary, then it is possible from the said GIRs, in principle, to derive many more particular invariance relations reasonable for using to solve various specific boundary-value problems of the RTT. Just the common character of such GIRs enables one to select various, convenient for studying, pairs of the solutions to boundary-value problems for the SRTE. In this respect, it is possible to get the information on these solutions in analytical, semi-analytical or other forms using the GIRs and their consequences. It should be especially underlined that the number of GIRs and their consequences is generally unlimited in principle due to the presence of an infinite set of algebraic operations (see the GIP) that remain radiation fields invariant almost everywhere. This aspect was not realized and was ignored in radiative transfer theory for a long time. Therefore, there were stated a whole number of the invariance principles (IPs) for plane-parallel media in the RTT (see, for example, publications (Ambartsumian, 1943a, 1943b; Chandrasekhar, 1950; van de Hulst, 1963; Shimizu, 1968; Hansen, 1969; Hunt and Grant, 1969a; Yengibarian and Mnatsakanian, 1974; Ivanov, 1975; Ivanov and Volkov, 1978; Yanovitskij, 1979, 1981; Viik, 1982 and references therein) on the basis, often, of the practically similar algebraic operations and actions (their algebraic nature was first simplified in the general form by Rogovtsov (1981a, 1981b)). It has already been mentioned that

such IPs are the particular cases of the GIP or its consequences. It will be useful to underline that the conclusions, which have been reached in the majority of these publications and which have been usually named as the invariance principles, are essentially statements on the existence of expressions of a balanced kind that connect together the solutions to different or related problems (these relations are, in fact, the invariance ones in their essence). The said IPs established the existence of partitions or representations of radiation field characteristics in plane-parallel media (sometimes in spherically symmetrical media) as certain unions (these unions were made by addition, subtraction, multiplication, and integration operations, etc.), the original elements of these unions being the solutions to boundary-value problems of the RTT for plane-parallel (or spherically symmetrical) media. In other words, these IPs as well as the GIP refer to a kind of statements that cannot be treated within the scope of the general theoretical group approach widely used in theoretical physics, mathematics, etc. (a more detailed explanation of this topic is given in Rogovtsov (1999)).

Now let a number of important particular cases of the general invariance relation of Eq. (7.16) be stated. The integration in Eq. (7.16) should be generally made over the whole unit sphere Ω. For solving a number of specific boundary-value problems of the RTT, it is useful to utilize the general invariance relations, in which the integration is made over hemispheres $\Omega_{\pm}$.

Let S_β be an absorber, i.e. the boundary of body V_β totally absorbs radiation incident on it. Note that such a situation occurs, for example, when body V_β is a non-concave one to be surrounded by the vacuum or by a purely absorbing medium and S_β is a totally radiation-transparent boundary. In this case, the GIR of Eq. (7.16) takes the form of the invariance relation of the 'first intersection' of the following general kind (Rogovtsov, 1981c, 1989, 1999):

$$
\begin{aligned}
&\Theta_{\mathrm{V}^0_\beta}\left(\vec{r}\right) I\left(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta\right)\\
&= \iiint\limits_{\mathrm{V}^0_\beta} d\mathrm{V}' \int\limits_{\Omega} d\Omega' \int\limits_{t_0-0}^{t} G_*\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}';t-t';\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right) g_\beta\left(\vec{r}',\vec{\Omega}',t';\mathrm{V}^0_\beta\right) dt'\\
&- \iint\limits_{\mathrm{S}_\beta(\mathrm{V}^0_\beta)} dS' \int\limits_{\Omega_+} \left(\vec{n}'\cdot\vec{\Omega}'\right) d\Omega' \int\limits_{t_0-0}^{t} G_*\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}';t-t';\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right) I\left(\vec{r}',\overrightarrow{\Omega'},t';\mathrm{V}_\beta\right) dt'\\
&+ \iiint\limits_{\mathrm{V}^0_\beta} d\mathrm{V}' \int\limits_{\Omega} G_*\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}';t-t_0;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right) \left(v\left(\vec{r}'\right)\right)^{-1} I\left(\vec{r}',\vec{\Omega}',t_0-0;\mathrm{V}_\beta\right) d\Omega',\\
&t > t_0.
\end{aligned}
\tag{7.17}
$$

Let boundary S_β partition body V_β on two bodies V_{β^*} and $\mathrm{V}_{\beta^{**}}$ ($\mathrm{V}_\beta = \mathrm{V}_{\beta^*} \cup \mathrm{V}_{\beta^{**}}$; symbol $\cup$ has the sense of the operation of the union of two sets). Denote the common boundary of these bodies by $\mathrm{S}_{\beta^*,\beta^{**}}$ $\left(\mathrm{S}_{\beta^*,\beta^{**}} \subset \mathrm{S}_\beta\right)$ and select V_{β_4} and S_{β_2} so that $\mathrm{V}^0_{\beta_4} = \mathrm{V}^0_\beta$ and surface $\mathrm{S}^{\times} = \mathrm{S}_{\beta_2} \backslash \mathrm{S}_{\beta^*,\beta^{**}}$ is an absorber. Then the GIR of Eq. (7.17) is an analog of the invariance relation of the 'first intersection' from the paper (Pikichian, 1982a).

If one takes an absorber as S_{β_2} in Eq. (7.16) and requires the fulfillment of the condition of $V^0_{\beta_4} = V^0_\beta$, then the invariance relation of the 'last intersection' is obtained in the general form. This particular case of the GIR of Eq. (7.16) has the form (Rogovtsov, 1981c, 1989, 1999):

$$
\begin{aligned}
& I\left(\vec{r},\vec{\Omega},t;V_\beta\right) \\
& = \iiint\limits_{V^0_\beta} dV' \int\limits_{\Omega} d\Omega' \int\limits_{t_0-0}^{t} G_*\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}';t-t';S_{\beta_2};V_{\beta_4}\right) g_\beta\left(\vec{r}',\vec{\Omega}',t';V^0_\beta\right) dt' \\
& - \iint\limits_{S_\beta(V^0_\beta)} dS' \int\limits_{\Omega_-} \left(\vec{n}'\cdot\vec{\Omega}'\right) d\Omega' \int\limits_{t_0-0}^{t} G_*\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}';t-t';S_{\beta_2};V_{\beta_4}\right) I\left(\vec{r}',\overrightarrow{\Omega'},t';V_\beta\right) dt' \\
& + \iiint\limits_{V^0_\beta} dV' \int\limits_{\Omega} G_*\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}';t-t_0;S_{\beta_2};V_{\beta_4}\right) \left(v\left(\vec{r}'\right)\right)^{-1} I\left(\vec{r}',\vec{\Omega}',t_0-0;V_\beta\right) d\Omega', \\
& t > t_0,
\end{aligned}
\tag{7.18}
$$

where $\vec{r}$ specifies points in V^0_β (but these points should not lie on underlying surfaces). If body V_β is represented as a union of bodies V_{β^*} and $V_{\beta^{**}}$ (i.e. $V_\beta = V_{\beta^*} \cup V_{\beta^{**}}$) and the common part of the boundaries of these bodies belongs to S_β, then the GIR of Eq. (7.18) is a analog of the invariance relations of the 'last intersection' from the paper (Pikichian, 1982a).

For the case of plane-parallel media and stationary sources, the invariance relations of the 'first and last intersections' were analyzed in detail by Ivanov (1976).

The generalization of the invariance relations of Eqs (7.17) and (7.18) is the invariance relations of the mixed type, in which, at one portion $S^\times_\beta(V^0_\beta)$ of set $S_\beta(V^0_\beta)$, the integration is made over the hemispheres of the Ω_+ kind and, at the other portion $S^{\times\times}_\beta(V^0_\beta)$, it is made over the hemispheres of the Ω_- kind (at this, $S^\times_\beta(V^0_\beta) \cap S^{\times\times}_\beta(V^0_\beta)$ is the curve or the empty set). The GIRs of such a type were first derived in the paper (Rogovtsov, 1981c).

Let body V_β be really a part of body V_{β_4} and S_{β_2} possesses the same optical characteristics as those of S_β (it follows from this, in particular, that the properties of S_β do not depend on the radiation field). Also let $g_\beta\left(\vec{r}',\vec{\Omega}',t';V^0_\beta\right) = \delta\left(\vec{r}'-\vec{r}_1\right)\delta\left(\vec{\Omega}'-\vec{\Omega}_1\right)\delta\left(t\right)$. Under the said conditions, the GIR of Eq. (7.16) assumes the following form (Rogovtsov , 1989):

$$
\begin{aligned}
& \Theta_{V^0_\beta}\left(\vec{r}\right) G_*\left(\vec{r};\vec{\Omega};\vec{r}_1,\vec{\Omega}_1,t;S_{\beta_2};V_{\beta_4}\right) = \Theta_{V^0_\beta}\left(\vec{r}_1\right) G_*\left(\vec{r};\vec{\Omega};\vec{r}_1,\vec{\Omega}_1,t;S_{\beta_2};V_{\beta_4}\right) \\
& \quad - \iint\limits_{S_\beta\left(V^0_\beta\right)} dS' \int\limits_{\Omega} \left(\vec{n}'\cdot\vec{\Omega}'\right) d\Omega' \int\limits_{-0}^{t} G_*\left(\vec{r};\vec{\Omega};\vec{r}',\vec{\Omega}',t-t';S_{\beta_2};V_{\beta_4}\right) \\
& \quad \times G_*\left(\vec{r}',\vec{\Omega}',\vec{r}_1,\vec{\Omega}_1,t';S_{\beta_2};V^0_\beta\right) dt', \ t > 0,
\end{aligned}
\tag{7.19}
$$

where $t > 0$ and radius vectors $\vec{r},\vec{r}_1$ specify points outside the boundary S_β. Note here that the GIRs of the type of Eq. (7.19) enable one to find the normalization

constants of the asymptotic formulae for the Green functions of Eq. (7.1) for the case of turbid media of various configurations (see, for example, Rogovtsov, 1986a, 1986c, 1988a, 1988c, 1988d, 1990a, 1990b, 1991a, 1991c, 1992a).

There were derived a number of other GIRs in the publications (Rogovtsov, 1980, 1981a, 1983, 1985b, 1986b, 1989, 1993, 1999; Pikichian, 1982a, 1982b, 1983, 1989), which, in principle, are the consequences of the GIR of Eq. (7.16). Some of these consequences will be given in subsection 7.2.4. Beside such types of the GIRs, there were derived the GIRs, which comprise functions $\Phi\bigl(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta\bigr)$, $\Phi\bigl(\vec{r},\vec{\Omega};\mathrm{V}_\beta\bigr)$ (or the source functions (Bekefi, 1966; Sobolev, 1975)) or the truncated Green functions for Eq. (7.1) (the general definitions of these functions are given by Rogovtsov (1980, 1981a).

7.2.3.2 General invariance relations for the case of stationary sources

Let internal and external sources are independent of time. The other assumptions that have been used to derive the GIRs of Eqs (7.16)–(7.19) are believed to be valid. Write down the stationary analogs of these GIRs.

The stationary analogs of the GIRs of Eqs (7.16)–(7.19) are the following (Rogovtsov 1981a, 1989, 1999):

$$
\begin{aligned}
&\Theta_{\mathrm{V}^0_\beta}(\vec{r})\, I\bigl(\vec{r},\vec{\Omega};\mathrm{V}_\beta\bigr)\\
&= \iiint\limits_{\mathrm{V}^0_\beta} d\mathrm{V}' \int\limits_{\Omega} G_*\bigl(\vec{r},\vec{\Omega};\vec{r}',\vec{\Omega}';\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\bigr)\, g_\beta\bigl(\vec{r}',\vec{\Omega}';\mathrm{V}^0_\beta\bigr)\, d\Omega'\\
&\quad - \iint\limits_{\mathrm{S}_\beta(\mathrm{V}^0_\beta)} dS' \int\limits_{\Omega} \bigl(\vec{n}'\cdot\vec{\Omega}'\bigr) G_*\bigl(\vec{r},\vec{\Omega};\vec{r}',\vec{\Omega}';\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\bigr) I\bigl(\vec{r}',\vec{\Omega}';\mathrm{V}_\beta\bigr)\, d\Omega'; \qquad (7.20)
\end{aligned}
$$

$$
\begin{aligned}
&\Theta_{\mathrm{V}^0_\beta}(\vec{r})\, I\bigl(\vec{r},\vec{\Omega};\mathrm{V}_\beta\bigr)\\
&= \iiint\limits_{\mathrm{V}^0_\beta} d\mathrm{V}' \int\limits_{\Omega} G_*\bigl(\vec{r},\vec{\Omega};\vec{r}',\vec{\Omega}';\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\bigr) g_\beta\bigl(\vec{r}',\vec{\Omega}';\mathrm{V}^0_\beta\bigr)\, d\Omega'\\
&\quad - \iint\limits_{\mathrm{S}_\beta(\mathrm{V}^0_\beta)} dS' \int\limits_{\Omega_+} \bigl(\vec{n}'\cdot\vec{\Omega}'\bigr) G_*\bigl(\vec{r},\vec{\Omega};\vec{r}',\vec{\Omega}';\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\bigr) I\bigl(\vec{r}',\vec{\Omega}';\mathrm{V}_\beta\bigr)\, d\Omega'; \qquad (7.21)
\end{aligned}
$$

$$
\begin{aligned}
&I\bigl(\vec{r},\vec{\Omega};\mathrm{V}_\beta\bigr)\\
&= \iiint\limits_{\mathrm{V}^0_\beta} d\mathrm{V}' \int\limits_{\Omega} G_*\bigl(\vec{r},\vec{\Omega};\vec{r}',\vec{\Omega}';\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\bigr) g_\beta\bigl(\vec{r}',\vec{\Omega}';\mathrm{V}^0_\beta\bigr)\, d\Omega'\\
&\quad - \iint\limits_{\mathrm{S}_\beta(\mathrm{V}^0_\beta)} dS' \int\limits_{\Omega_-} \bigl(\vec{n}'\cdot\vec{\Omega}'\bigr) G_*\bigl(\vec{r},\vec{\Omega};\vec{r}',\vec{\Omega}';\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\bigr) I\bigl(\vec{r}',\overrightarrow{\Omega'};\mathrm{V}_\beta\bigr)\, d\Omega'; \qquad (7.22)
\end{aligned}
$$

$$\Theta_{\mathrm{V}^0_\beta}(\vec{r})\, G_*\big(\vec{r},\vec{\Omega};\vec{r}_1,\vec{\Omega}_1;\mathrm{S}_\beta;\mathrm{V}_{\beta_4}\big) = \Theta_{\mathrm{V}^0_\beta}(\vec{r}_1)\, G_*\big(\vec{r},\vec{\Omega};\vec{r}_1,\vec{\Omega}_1;\mathrm{S}_\beta;\mathrm{V}_{\beta_4}\big)$$
$$- \iint\limits_{\mathrm{S}_\beta(\mathrm{V}^0_\beta)} dS' \int\limits_{\Omega} \big(\vec{n}'\cdot\vec{\Omega}'\big) G_*\big(\vec{r},\vec{\Omega};\vec{r}',\vec{\Omega}';\mathrm{S}_\beta;\mathrm{V}_{\beta_4}\big) G_*\big(\vec{r}',\vec{\Omega}';\vec{r}_1,\vec{\Omega}_1;\mathrm{S}_\beta;\mathrm{V}_{\beta_4}\big)\, d\Omega'. \tag{7.23}$$

Note that all the GIRs mentioned and written down above can be derived by means of formal mappings too, based on the usage of the invariant redefinition method and the Fourier transform theory. This was done by Rogovtsov and Samson (1976); Rogovtsov (1980, 1981a). Besides, the relationships between the said formal mappings and the GIRs were established in these papers.

Let us turn now to the fact that some particular versions of the GIRs for stationary sources can be formally derived by means of the surface pseudo-sources method (Beauwens and Devooght, 1968; Laletin, 1969). However, the pseudo-sources do not always have a physical sense. Note, among the most prominent results obtained before the statement of the GIP and the derivation of the GIRs, an important relation (using the GIP, it can be interpreted as the invariance relation) that comprises the volume Green functions of Eq. (7.10) for the case of infinite and finite media, when $m(\vec{r}) \equiv 1$. This relation was formally derived by Case (1969) directly from the fundamental identity. The identity was obtained by Case (1957) for $m(\vec{r}) \equiv \mathrm{const}$ by using Eq. (7.10) without attracting any physical considerations. The results obtained by Case (1969); Beauwens and Devooght (1968); Laletin (1969) are directly connected with Placzek's lemma proven in the book by Case, Hoffman, and Placzek (1953). There it is stated that the solution to Eq. (7.10) with $m(\vec{r}) \equiv 1$ inside some body V remains unchanged, if, together with the internal sources contained therein, it is 'imbedded into an infinite medium and if, simultaneously, corresponding 'pseudo sources' (they should be negative) will be introduced on the boundary of body V. None of these publications, however, considered media of complex configurations that can have arbitrary underlying surfaces (their portions, in particular, can be two- or one-sided ones). Besides, these publications provided no physical interpretation; they derive any analytical or semi-analytical solutions to boundary-value problems of the RTT for media that did not possess the plane-parallel symmetry.

7.2.4 Scheme of using the general invariance principle and the general invariance relations

While solving specific boundary-value problems of the RTT on the basis of the GIP, reasonable GIRs should be constructed by means of algebraic operations (see subsection 7.2.2) and using the selection procedure of bodies of the kind V_β and V_{β_4} from set $\{\mathrm{V}_\beta\}_{\beta\in B}$ possessing required properties (the selection procedure is substantially similar to the redefinition and embedding procedures). These GIRs should relate a desired solution (or characteristics of the solution) to a specific boundary-value problem of the RTT for the case of bodies of the kind V_{β_4}. When searching for the required GIRs, it is helpful to utilize a number of features of the GIP and the general physical conceptions of the radiative transfer process. In particular, a convenient accessory means for finding the required GIRs is the partition

of photon 'trajectories', which randomly walk in media of bodies of the kind V_β and V_{β_4}, on topologically different classes. Moreover, it should be accounted for the relative character of the boundary conception for the same body and the possibility to change the conception while implementing the procedure of imbedding body V_β into body V_{β_4}. Such a situation is due to the presence of an infinite set $\mathfrak{R}$ of operations, with which the geometrical and physical sense imposed on the boundaries of bodies V_β and V_{β_4} can be changed. Note also that the above written down relations or relations derived in the said publications can be, for example, taken as the GIRs. Second, since the GIRs and their consequences include usually the characteristics of radiation fields in bodies of similar and different shapes, then, by changing the configuration of, at least, one body, it is possible to reduce the solution to complex boundary-value problems of the RTT to the investigation of simpler boundary-value problems of the RTT that can be solved, if only partially. This enables one to establish a certain hierarchy of boundary-value problems of the RTT and to start the solution from the very simplest ones. After the solution to these simple problems, it is possible, with using the GIRs, to come to the investigation of the properties of radiation fields in media of more and more complicated configurations. Third, on the base of any information on the properties of bodies of kind V_β and V_{β_4} and on some quantities (in particular, on the Green function) entering the GIRs, it is possible to derive more particular invariance relations. It is desirable for this that the relations themselves have direct physical and mathematical interest and can be utilized for getting the information on the desired solutions to boundary-value problems of the RTT in analytical, semi- analytical or other forms. The GIRs and these particular invariance relations as well as the derivation of particular consequences from them should be evaluated by accounting for the structure of the GIRs (in particular, when $\vec{r} = \vec{r}_{S_\beta} - 0\vec{n}$, many of the GIRs are the integral equations for the desired solutions to original boundary-value problems) and of the said particular invariance relations themselves. Besides, it is useful to take into account the possibility to isolate, from the Green functions entering the said relations, their principal parts describing analytically the main contribution to the radiation field in certain regions of bodies of the kind V_β and V_{β_4}. In particular, such situations occur in those regions of the above bodies, where the depth regimes are established. Fourth, while using the GIRs and their consequences, it is useful to put into attention the exact solutions or the strict inequalities and simultaneously to take into account the availability of various mathematical methods for solving the equations of different types. Moreover, the general physical information on the radiative transfer process within the frame of the original boundary-value problem of the RTT often enables the ways for getting analytical or semi-analytical information from the particular invariance relations and from their various modifications to be more effectively found. Among the advantages of an approach based on the GIP, there should be mentioned its common character, heuristicity, and sufficient flexibility and constructive nature, while solving boundary-value problems of the RTT for media of various configurations. The drawback of the same approach as well as of the majority of other general RTT methods is that it cannot provide the equivalent consideration of all the problems arising in the applications. However, the possibility of deriving the solutions (in analytical or semi- analytical forms) to the whole classes of multidimensional boundary-value problems of the RTT on

this basis points to its wide opportunities that supplement the capabilities of the numerical methods and the Monte Carlo method.

The evidence described in subsections 7.2.2 and 7.2.3 and the monograph (Rogovtsov, 1999) comprise the basis of the general invariance relations reduction method (GIRRM). A lot of specific examples of using the GIRRM for solving direct and inverse problems of the RTT have been given, in particular, by Rogovtsov and Samson (1976, 1985a, 1985b, 1987); Rogovtsov (1980, 1981a, 1981b, 1981c, 1983, 1985a, 1985b, 1985c, 1985d, 1986a, 1986b, 1986c, 1988a, 1988b, 1988c, 1988d, 1989, 1990a, 1990b, 1991a, 1991b, 1992a, 1992b, 1993, 1997, 1999, 2002, 2007); Anisimov and Rogovtsov (2002); Pikichian (1982a, 1982b, 1983, 1989); Rogovtsov and Karpuk (2003); Rogovtsov and Borovik (2009). A number of the results illustrating the capabilities of the GIRRM will be presented in the next subsections of this chapter.

7.3 Some general examples of using the general invariance relations reduction method

The above GIRs are of rather a general character and, therefore, allow various simplifications, modifications, and combinations with each other. By making one or another assumptions on the properties of turbid media (in particular, on their configurations), it is possible, on the basis of these GIRs, to get a lot of various particular versions that will reduce the original boundary-value problem of the RTT to the solution to simpler problems. Moreover, such relations can connect together physically important characteristics of radiation fields in turbid media of complex configurations to enable the analytical expressions and estimations to be obtained for various quantities. This section will provide a number of such examples.

7.3.1 Doubling formulae

Consider a macroscopically homogeneous non-concave medium V_β, which boundary S_β is totally transparent for radiation (i.e. $\hat{R} = \hat{0}$ and $\hat{T} = \hat{E}$). Let body V_β be such that, after the section of this body by surface $\Xi_j\big(\vec{b}\big)$ that is obtained from S_β with the translation at any vector $\vec{b}$ inside the turbid medium, one is able to isolate, from V_β, a part $\mathrm{V}_{\beta^*}\big(\vec{b}\big)$ of body V_β that is identical to the initial body V_β by its shape. As an example of such a body, indicate a body having the shape of a semi-infinite cylindrical region bounded by a non-concave surface. One gets for the surface Green function $G_{\mathrm{S}_\beta}\big(\vec{r};\vec{\Omega};\vec{r}_{\mathrm{S}_\beta} - 0\vec{n},\vec{\Omega}_1,t;\mathrm{V}_\beta\big)$ of Eq. (7.1) the formula from the GIR of Eq. (7.18) with accounting for the properties of body V_β (Rogovtsov, 1981c):

$$
\begin{aligned}
&G_{\mathrm{S}_\beta}\big(\vec{r},\vec{\Omega};\vec{r}_{\mathrm{S}_\beta} - 0\vec{n},\vec{\Omega}_1;t;\mathrm{V}_\beta\big)\\
&= \iint\limits_{\mathrm{S}_\beta(\mathrm{V}^0_\beta)} dS' \int\limits_{\Omega_-} d\Omega' \int\limits_{-0}^{t} G_{\mathrm{S}_\beta}\big(\vec{r}-\vec{b},\vec{\Omega};\vec{r}',\vec{\Omega}',t-t';\mathrm{V}_\beta\big)\\
&\quad\times G_{\mathrm{S}_\beta}\big(\vec{r}'+\vec{b},\vec{\Omega}';\vec{r}_{\mathrm{S}_\beta} - 0\vec{n},\vec{\Omega}_1;t';\mathrm{V}_\beta\big)\,dt',
\end{aligned}
\tag{7.24}
$$

where $t > 0$, $\vec{r}$ specifies points inside body $\mathrm{V}_{\beta^*}\big(\vec{b}\big)$, and $\vec{n} \cdot \vec{\Omega}_1 < 0$ $\big(\vec{n} = \vec{n}\,\big(\vec{r}_{\mathrm{S}_\beta}\big)\big)$. One can replace $\vec{r}_{\mathrm{S}_\beta} - 0\vec{n}$ in Eq. (7.24) by $\vec{r}_{\mathrm{S}_\beta}$, because body V_β does not comprise any underlying surfaces. If one sets $\vec{r} = \vec{r}_{\mathrm{S}_\beta} + 2\vec{b}$ in Eq. (7.24), then we will obtain the below doubling formula for $t > 0$ and $\vec{n} \cdot \vec{\Omega}_1 < 0$:

$$
\begin{aligned}
&G_{\mathrm{S}_\beta}\big(\vec{r}_{\mathrm{S}_\beta} + 2\vec{b}, \vec{\Omega}; \vec{r}_{\mathrm{S}}, \vec{\Omega}_1, t; \mathrm{V}_\beta\big) \\
&= \iint\limits_{\mathrm{S}_\beta(\mathrm{V}^0_\beta)} dS' \int\limits_{\Omega_-} d\Omega' \int\limits_{-0}^{t} G_{\mathrm{S}_\beta}\big(\vec{r}_{\mathrm{S}_\beta} + \vec{b}, \vec{\Omega}; \vec{r}', \vec{\Omega}'; t - t'; \mathrm{V}_\beta\big) \\
&\quad \times G_{\mathrm{S}_\beta}\big(\vec{r}' + \vec{b}, \vec{\Omega}'; \vec{r}_{\mathrm{S}_\beta}, \vec{\Omega}_1; t'; \mathrm{V}_\beta\big)\, dt', \qquad (7.25)
\end{aligned}
$$

This formula is the generalization of the known doubling formulae (see, for example, Yanovitskij, 1997) derived for plane-parallel turbid media irradiated by stationary, infinitely wide, mono-directional radiation sources. The formula enables one to reduce the computation time, while searching for $G_{\mathrm{S}_\beta}(\dots)$ for bodies of the said shape. It is also useful for deriving the asymptotic formulae for the surface Green functions under depth regimes.

Let turbid medium V_β be not irradiated by external radiation and be imbedded in 'vacuum' or in a purely absorbing medium. Assume that boundary S_β of body V_β is non-concave and totally transparent for radiation. Besides, we will believe that there are no underlying surfaces inside V_β (i.e. inside V^0_β), but there is a radiation 'source' with density $\delta\,(\vec{r} - \vec{r}_1)\,\delta\big(\vec{\Omega} - \vec{\Omega}_1\big)\delta\,(t)$ (the word 'source' is put in inverted commas, since the density has no dimensional density). Under the above conditions, it follows from the GIR of Eq. (7.16) (Rogovtsov, 1983):

$$
\begin{aligned}
&G_*\big(\vec{r}, \vec{\Omega}; \vec{r}_1, \vec{\Omega}_1; t_1 + t_2; \mathrm{S}_\beta; \mathrm{V}_\beta\big) \\
&= \iiint\limits_{\mathrm{V}^0_\beta} d\mathrm{V}' \int\limits_{\Omega} G_*\big(\vec{r}, \vec{\Omega}; \vec{r}', \vec{\Omega}'; t_1; \mathrm{S}_\beta; \mathrm{V}_\beta\big) \\
&\quad \times G_*\big(\vec{r}', \vec{\Omega}'; \vec{r}_1, \vec{\Omega}_1; t_2; \mathrm{S}_\beta; \mathrm{V}_\beta\big)\,(\upsilon\,(\vec{r}'))^{-1}\, d\Omega', \qquad (7.26)
\end{aligned}
$$

where $\vec{r}$ specifies points in V^0_β, and $t_1 > 0$, $t_2 > 0$. Eq. (7.26) enables the volume Green function at time moment $(t_1 + t_2)$ to be found, if it is known in time moments t_1 and t_2 at any points lying in V^0_β. It is helpful to use Eq. (7.26) for the derivation of the normalization constants of the asymptotic formulae for function $G_*\big(\vec{r}, \vec{\Omega}; \vec{r}_1, \vec{\Omega}_1; t; \mathrm{S}_\beta; \mathrm{V}_\beta\big)$ at $t \to +\infty$. If $t_1 = t_2$, then Eq. (7.26) is the doubling formula for the case of non-stationary sources.

7.3.2 On the relationship between the volume Green functions and the generalized reflection function

Let there be a 'source' with density $\delta\,(\vec{r} - \vec{r}_1)\,\delta\big(\vec{\Omega} - \vec{\Omega}_1\big)\delta\,(t)$ in a turbid medium V_β, which is of a non-concave shape and does not comprise any underlying surfaces. We will assume that the surrounding of body V_β is a purely absorbing medium or vacuum that does not affect the radiation field in V_β. Take a medium, which also does not comprise any underlying surfaces and is of a non-concave shape, as body

V_{β_4}. It is really assumed that there are no discontinuities of the refractive index in V_β and V_{β_4}. Since no essential limitations on the spatial dependences of other local optical characteristics of V_β and V_{β_4} are imposed, then the assumption on their non-concave character does not limit the commonality of the consideration.

Using Eq. (7.1) (or Eq. (7.5)), the properties of the Dirac δ-function, and the classical principle of the optical reciprocity (Case and Zweifel, 1967) for the case of media with $m(\vec{r}) \equiv \text{const}$, one can show that the following equality is valid:

$$m_1(\vec{r}_1, \vec{r})\, G_*\big(\vec{r}, \vec{\Omega}; \vec{r}_1, \vec{\Omega}_1, t; S_\beta; V_\beta\big) = m_1(\vec{r}, \vec{r}_1)\, G_*\big(\vec{r}_1, -\vec{\Omega}_1; \vec{r}, -\vec{\Omega}; t; S_\beta; V_\beta\big). \tag{7.27}$$

If the refractive index of turbid medium V_β is constant, then Eq. (7.27) takes the form of the classical principle of the optical reciprocity.

Within the scope of the assumptions made, by twofold use of the GIR of Eq. (7.16) and the expression of Eq. (7.27), it is possible to verify the validity of the following formulae:

$$\begin{aligned} G_*\big(\vec{r}, \vec{\Omega}; \vec{r}_1, \vec{\Omega}_1; t; S_\beta; V_\beta\big) &= G_*\big(\vec{r}; \vec{\Omega}; \vec{r}_1, \vec{\Omega}_1, t; S_\beta; V_{\beta_4}\big) \\ &- \iint\limits_{S_\beta(V_\beta^0)} dS' \int\limits_{\Omega_+} \big(\vec{n} \cdot \vec{\Omega}'\big)\, d\Omega' \int\limits_{-0}^{t} G_*\big(\vec{r}; \vec{\Omega}; \vec{r}', \vec{\Omega}', t - t'; S_\beta; V_{\beta_4}\big) \\ &\times G_*\big(\vec{r}, \vec{\Omega}; \vec{r}_1, \vec{\Omega}_1, t'; S_\beta; V_\beta\big)\, dt', \end{aligned} \tag{7.28}$$

$$\begin{aligned} G_*\big(\vec{r}'_{S_\beta}, \vec{\Omega}'; \vec{r}_1, \vec{\Omega}_1, t; S_\beta; V_\beta\big) &= G_*\big(\vec{r}'_{S_\beta}, \vec{\Omega}'; \vec{r}_1, \vec{\Omega}_1, t; S_\beta; V_{\beta_4}\big) \\ &- \frac{1}{\pi} \iint\limits_{S_\beta(V_\beta^0)} dS'' \int\limits_{\Omega_+} \big(\vec{n}'' \cdot \vec{\Omega}''\big)\, d\Omega'' \int\limits_{-0}^{t} G_*\big(\vec{r}'', -\vec{\Omega}''; \vec{r}_1, \vec{\Omega}_1, t - t'; S_\beta; V_{\beta_4}\big) \\ &\times R\big(\vec{r}'_{S_\beta}, \vec{\Omega}'; \vec{r}'', -\vec{\Omega}''; t'; S_\beta; V_\beta\big)\, dt'. \end{aligned} \tag{7.29}$$

Here $R\big(\vec{r}'_{S_\beta}, \vec{\Omega}'; \vec{r}''_{S_\beta}, -\vec{\Omega}''; t; S_\beta; V_\beta\big) = \pi\big(\big|\vec{n}'' \cdot \vec{\Omega}''\big|\big)^{-1} G_{S_\beta}\big(\vec{r}'_{S_\beta}, \vec{\Omega}'; \vec{r}''_{S_\beta}, -\vec{\Omega}'', t; S_\beta; V_\beta\big)$ is the generalized reflection function; radius-vector $\vec{r}$ specifies points in V_β^0. The generalized reflection function characterizes radiation exiting body V_β, when an infinitely narrow mono-directional beam of external radiation is incident on boundary S_β.

It follows from Eqs (7.28) and (7.29) that it is sufficient to know the volume Green function for Eq. (7.1) for the case of body V_{β_4} and the generalized reflection function $R(\dots)$ to find the volume Green function for Eq. (7.1) for the case of body V_β. The formulae of Eqs (7.28) and (7.29) are most interesting for situations, when V_β and V_{β_4} are the macroscopically homogeneous turbid bodies and when infinite turbid medium V_∞ is being taken as V_{β_4}. The infinite turbid medium V_∞ is understood everywhere below as a medium, for which integral $\int_0^{+\infty} \kappa_{ext}\big(\vec{r} - \xi\vec{\Omega}\big)\, d\xi$ diverges for any $\vec{\Omega} \in \Omega$ and $\vec{r}$. Equations (7.28) and (7.29) coincide with the relations derived earlier by Rogovtsov (1981c) for the particular case, when $m(\vec{r}) \equiv \text{const}$ in V_β. Some stationary analogs of these relations have been considered in detail by Rogovtsov and Samson (1985a, 1985b).

7.3.3 Analog of the Kirchhoff law for the case of non-equilibrium radiation in turbid media

Let V_β be a turbid medium of an arbitrary shape. Assume that all the real surrounding of body V_β is a turbid medium or a medium that does not scatter radiation at all. We will believe that there are no discontinuities of the refractive index in body V_β and in its real surrounding (i.e. there are no underlying surfaces). Besides, assume that body V_β and its real surrounding can be positioned in an infinite or finite part E_0^3 of space E^3, the part E_0^3 is non-concave and is bounded by surface Ξ^*. Additional to the assumptions made, it is required that body V_β and its real surrounding can be embedded in infinite medium V_∞ without changing the characteristics of radiation fields in them. Now, it is believed that the scalar radiative transfer theory is applicable to body V_β, its surrounding, and V_∞. Such an embedding can be, for example, implemented by positioning an absorbing medium without any radiation sources in a part $(E^3 \backslash E_0^3)$ of space E^3. Fig. 7.4 shows one of the possible versions to position turbid medium V_β and its real surrounding (in this example, body V_β is disconnected, and $V_\beta = V_{\beta'} \cup V_{\beta''} \cup V_{\beta'''}$, i.e. V_β is the union of set $V_{\beta'}$, $V_{\beta''}$, and $V_{\beta'''}$). All the space E^3 illustrated in Fig. 7.4 is filled by medium V_∞. The latter is the union of body V_β, its surrounding located in $(E_0^3 \backslash V_\beta)$, and an absorbing medium (difference $E_0^3 \backslash V_\beta$ of sets E_0^3 and V_β is understood in the geometrical sense).

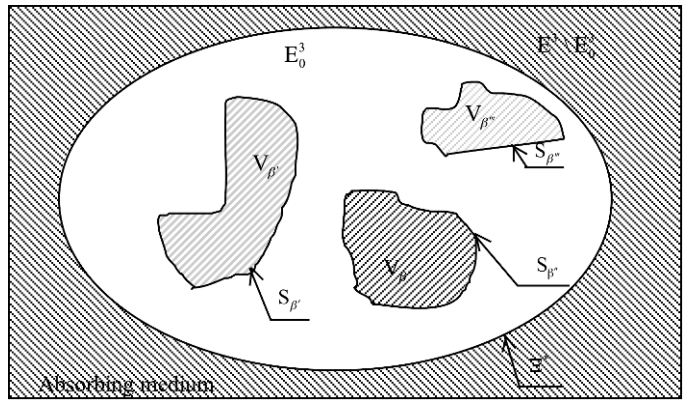


Fig. 7.4. Example of positioning body V_β, its surrounding, and absorbing medium outside of non-concave surface Ξ^* bounding part E_0^3 of space E^3.

To derive the desired analog of the Kirchhoff law and a number of other results, we require the exact solutions to the SRTE of Eq. (7.1) (or of Eq. (7.5) and Eq. (7.6)) that are written down for infinite medium V_∞. Assume that in such a medium there are distributed the radiation sources with density $g\left(\vec{r}, \vec{\Omega}, t; V_\infty^0\right) = \left[\left(\kappa_{ext}\left(\vec{r}\right) - \kappa_{sca}\left(\vec{r}\right)\right)\varphi\left(t\right) + \left(v\left(\vec{r}\right)\right)^{-1}\varphi'(t)\right] m^2\left(\vec{r}\right)$, where function $\varphi(t)$ is nonnegative and is equal to zero for $t < \varepsilon$ (here ε is an infinitesimal positive number). Let the initial radiation field be absent in V_∞. i.e. $I\left(\vec{r}, \vec{\Omega}, -0; V_\infty\right) \equiv 0$. Now account for that function $\bar{I}_{ext}(\ldots)$ in Eq. (7.6) should be assumed to be identically zero

for the case of medium V_∞. Besides, let Eq. (7.6) have a single solution under the condition of $\int_0^{+\infty} \kappa_{ext}\big(\vec{r}-\xi\vec{\Omega}\big)\,d\xi = +\infty$ for any $\vec{r}$ and $\vec{\Omega}$ (this condition is fulfilled for medium V_∞). Within the frame of the made assumption, it is easy to be convinced that the exact solution to Eq. (7.6) has the following form:

$$\bar{\Phi}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\infty\big) = \bar{\varphi}\,(\sigma)\left(\kappa_{ext}\,(\vec{r}) + \sigma\,(v(\vec{r}))^{-1}\right)(m\,(\vec{r}))^2\,. \tag{7.30}$$

In its own turn, using the formulae of Eqs (7.7) and (7.30), it can be shown that under the above conditions the exact solution to Eq. (7.5) is

$$\bar{\mathrm{I}}\big(\vec{r},\vec{\Omega},\sigma;\mathrm{V}_\infty\big) = (m\,(\vec{r}))^2\,\bar{\varphi}\,(\sigma)\,. \tag{7.31}$$

The normalization condition for the phase function

$$(4\pi)^{-1}\int\limits_{\Omega} p\big(\vec{r};\Omega\cdot\vec{\Omega}'\big)\,d\Omega' = \frac{1}{2}\int\limits_{-1}^{1} p\,(\vec{r};\mu)\,d\mu \equiv 1$$

was taken into account, while deriving Eqs (7.30) and (7.31). It follows from Eqs (7.30) and (7.31) that the following equalities are valid:

$$\bar{\Phi}\big(\vec{r},\vec{\Omega},t;\mathrm{V}_\infty\big) = \big(\kappa_{ext}\,(\vec{r})\,\varphi\,(t) + (v(\vec{r}))^{-1}\,\varphi'\,(t)\big)\,(m\,(\vec{r}))^2\,, \tag{7.32}$$

$$\bar{\mathrm{I}}\big(\vec{r},\vec{\Omega},t;\mathrm{V}_\infty\big) = \big(m^2\,(\vec{r})\big)^2\,\varphi\,(t)\,, \quad t \geq -\varepsilon. \tag{7.33}$$

Let all the above-mentioned assumptions be satisfied, and let there be a point mono-directional source with density $\delta\,(\vec{r}-\vec{r}_1)\,\delta\big(\vec{\Omega}-\vec{\Omega}_1\big)\varphi\,(t)$ inside V_∞ (i.e. in V_β^0) or in region $(\mathrm{E}_0^3\backslash\mathrm{V}_\beta)$. Now use the GIR of Eq. (7.19), when an infinite medium V_∞ is taken as body V_{β_4}. Apply first the Laplace transform to this GIR and multiply the result by function $(m\,(\vec{r}))^{-2}\left(\kappa_{ext}(\vec{r}) + \sigma\,(v\,(\vec{r}))^{-1} - \kappa_{sca}(\vec{r})\right)\gamma\bar{\varphi}(\sigma)$, where γ is the constant having the dimension inversely proportional to the product of distance squared by time. Then integrate the obtained expression over variables $\vec{\Omega}$ and $\vec{r}$ of sets Ω and V_∞, respectively. If one takes into account Eq. (7.27), the definition of the volume Green function for Eq. (7.1), and Eqs (7.30)–(7.33), then the following analog of the Kirchhoff law is deduced from the relation derived by the above-said way:

$$\bar{\mathrm{I}}_1\big(\vec{r}_1,-\vec{\Omega}_1,\sigma;\mathrm{V}_{\beta^*}\big)\big(\alpha_1\big(\vec{r}_1,\vec{\Omega}_1,\sigma;\mathrm{V}_{\beta^*}\big)\big)^{-1} = \gamma\bar{\varphi}\,(\sigma)\,, \tag{7.34}$$

where

$$\begin{aligned}
&\bar{\mathrm{I}}_1\big(\vec{r}_1,-\vec{\Omega}_1,\sigma;\mathrm{V}_{\beta^*}\big)\\
&= \gamma\bar{\varphi}\,(\sigma)\iiint\limits_{\mathrm{V}_\beta^0} d\mathrm{V}\int\limits_{\Omega}\bar{\mathrm{G}}_*^1\big(\vec{r}_1,-\vec{\Omega}_1,\vec{r},\vec{\Omega};\sigma;\mathrm{S}_\beta;\mathrm{V}_\infty^*\big)\big(\kappa_{ext}^*\,(\vec{r};\sigma) - \kappa_{sca}(\vec{r})\big)\,d\Omega\\
&= \gamma\bar{\varphi}\,(\sigma)\,\bar{\mathrm{I}}^{\times}\big(\vec{r}_1,-\vec{\Omega}_1;\sigma;\mathrm{S}_\beta;\mathrm{V}_\infty^*\big);
\end{aligned} \tag{7.35}$$

$$\alpha_1\left(\vec{r}_1,\vec{\Omega}_1;\sigma;\mathrm{V}_{\beta^*}\right)$$
$$= \Theta_{\mathrm{V}^0_\beta}(\vec{r}) - \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right)} ds' \int\limits_{\Omega} \left(\vec{n}'\cdot\vec{\Omega}'\right)\bar{\mathrm{G}}^1_*\left(\vec{r}',\vec{\Omega}';\vec{r}_1,\vec{\Omega}_1;\sigma;\mathrm{S}_\beta;\mathrm{V}^*_\infty\right) d\Omega'. \quad (7.36)$$

Function $\bar{\mathrm{G}}^1_*\left(\vec{r},\vec{\Omega};\vec{r}_1,\vec{\Omega}_1;\sigma;\mathrm{S}_\beta;\mathrm{V}^*_\infty\right)$ of Eq. (7.35) stands for the Laplace image of the volume Green function for Eq. (7.1), where the identity of $m(\vec{r}) \equiv \text{const}$ is valid for any $\vec{r}$. Here, however, the real velocity of light in a medium (i.e. in body V_β and outside of it) should be understood as a function $v(\vec{r})$. Function $\bar{\mathrm{G}}^1_*(\dots)$ satisfies Eq. (7.5), in which one should assume that $m(\vec{r}) \equiv \text{const}$ and $\mathrm{I}\left(\vec{r},\vec{\Omega},-0;\mathrm{V}_\beta\right) \equiv 0$ and replace $\bar{g}\left(\vec{r},\vec{\Omega},\sigma;\mathrm{V}^0_\beta\right)$ by$\delta\left(\vec{r}-\vec{r}_1\right)\delta\left(\vec{\Omega}-\vec{\Omega}_1\right)$. It should be noted that function $\sigma\left(v(\vec{r})\right)^{-1}$ entering Eq. (7.5) as a term of $\kappa^*_{ext}(\vec{r};\sigma)$ has the sense of temporal absorption. This term was explained in detail by Case and Zweifel (1967). In its turn, symbol V^*_∞ is understood as a model turbid medium, which has the same local optical characteristics as V_∞, except for the refractive index. If turbid medium V_∞ is macroscopically homogeneous and $\sigma = 0$, then $\mathrm{V}^*_\infty = \mathrm{V}_\infty$ from the viewpoint of local optical properties. Now let constant γ be selected, so that quantity $\gamma\bar{\varphi}(\sigma)$ coincides with the radiation intensity of the absolutely black body. Then functions $\bar{\mathrm{I}}_1\left(\bar{r}_1,-\bar{\Omega}_1,\sigma;\mathrm{V}_{\beta^*}\right)$ and $\alpha_1\left(\vec{r}_1,\vec{\Omega}_1;\sigma;\mathrm{V}_{\beta^*}\right)$ with nonnegative parameter σ can be interpreted as the radiative and absorptive abilities of some model turbid medium V_{β^*} (it is a portion of V^*_∞). The complete analogy of Eq. (7.34) with the Kirchhoff law occurs when refractive index $m(\vec{r}) \equiv const$ and $\sigma = 0$ in the medium V_∞. It is also useful to note that there is a simple relationship between the volume Green functions $\bar{\mathrm{G}}^1_*\left(\vec{r}',\vec{\Omega}';\vec{r},\vec{\Omega};\sigma;\mathrm{S}_\beta;\mathrm{V}^*_\infty\right)$ and $\bar{\mathrm{G}}_*\left(\vec{r}',\vec{\Omega}';\vec{r},\vec{\Omega},\sigma;\mathrm{S}_\beta;\mathrm{V}_\infty\right)$, which correspond to media V^*_∞ and V_∞. It is expressed by the following formula (Rogovtsov, 1999):

$$\bar{\mathrm{G}}^1_*\left(\vec{r}',\vec{\Omega}';\vec{r},\vec{\Omega};\sigma;\mathrm{S}_\beta;\mathrm{V}^*_\infty\right) = m_1(\vec{r},\vec{r}')\,\bar{\mathrm{G}}_*\left(\vec{r}',\vec{\Omega}';\vec{r},\vec{\Omega},\sigma;\mathrm{S}_\beta;\mathrm{V}_\infty\right). \quad (7.37)$$

Various particular cases of Eq. (7.34) have been obtained earlier by Sobolev (1973a, 1973b) and Rogovtsov (1981a, 1989).

7.3.4 General invariance relations for monochromatic radiation fluxes

To find monochromatic and integral (i.e. integrated over the frequency) radiation fluxes through surfaces that are the boundaries of turbid media or lie inside them is an important problem of the radiative transfer theory. The radiation energy absorbed by the whole turbid medium or by its portion is, in particular, expressed via these fluxes, which we need to know, for example, to derive the spherical albedo of planetary atmospheres (Sobolev, 1975) and to solve a problem on radiative exchange in the cloudy atmosphere (Minin, 1974, 1988; Avaste et al., 1978; Romanova, 1985).

Consider a turbid medium V_β, whose boundary S_β can be an underlying surface. Assume that S_β includes all the underlying surfaces contained in V_β. Do not however, let any simple piece of surface S_β osculate with the interior portion V^0_β of body V_β with both its sides simultaneously. Fig. 7.5 shows examples of the admissible and inadmissible shapes of bodies V_β and types of its boundaries. Let

the values of the refractive index are independent of the position of an observation point on $S_\beta(V^0_\beta)$. Let there be in V^0_β distributed internal radiation sources specified by density $\Theta_{V^0_\beta}(\vec{r})\, g_\beta(\vec{r}, \vec{\Omega}, t; V^0_\beta)$. Body V_β can also be illuminated by external radiation and, at the initial time moment $t_0 = -0$, there can exist some radiation intensity distribution $\Theta_{V^0_\beta}(\vec{r})\, I(\vec{r}, \vec{\Omega}, -0; V_\beta)$ in V^0_β that will be believed to be specified. Assume also that turbid medium V_β is not an infinite one.

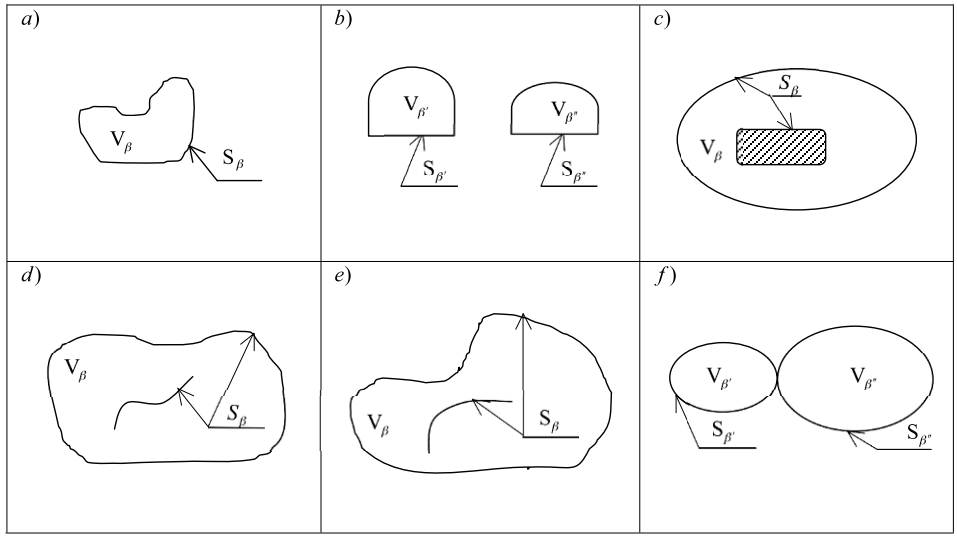


Fig. 7.5. Examples of admissible (a, b, c) and inadmissible (d, e, f) shapes of bodies V_β and types of its boundaries ($V_\beta = V_{\beta'} \cup V_{\beta''}$ and $S_\beta = S_{\beta'} \cup S_{\beta''}$).

Take the difference between the monochromatic power values (they correspond to a unit spectral range) that exit V^0_β and enter V^0_β through $S_\beta(V^0_\beta)$ at time moment t as a quantity being investigated. This quantity equals

$$F_1\left(S_\beta\left(V^0_\beta\right), t\right) = \iint\limits_{S_\beta\left(V^0_\beta\right)} dS \int\limits_{\Omega} \left(\vec{n} \cdot \vec{\Omega}\right) I(\vec{r}, \vec{\Omega}; t; V_\beta)\, d\Omega. \tag{7.38}$$

Quantity $F_1\left(S_\beta\left(V^0_\beta\right), t\right)$ is everywhere below named as a monochromatic flux through the boundary of body V_β or simply as a monochromatic radiance of body V_β. Now write down the GIR for $F_1\left(S_\beta\left(V^0_\beta\right), t\right)$. To do this, use the GIR of Eq. (7.16) at $t_0 = 0$ and the above-made assumptions. Then take a totally radiation-transparent surface (in the geometrical sense, it coincides with S_β) as S_{β_2} in the GIR of Eq. (7.16). Respectively, take a body, which has no underlying surfaces and contains a portion identical to the interior portion of body V_β (i.e. V^0_β) by its geometrical and optical properties, as body V_{β_4} in the GIR of Eq. (7.16). Assume also that body V_{β_4} is an infinite turbid medium or has a non-concave shape and can be embedded in V_∞.

Now introduce a surface S_{β^*} that lies in V^0_β and is obtained from S_β by using continuous deformation of the shape of the latter. Surface bounds some body V_{β^*} ($\mathrm{V}_{\beta^*} \subset \mathrm{V}^0_\beta$). If one first transforms the GIR of Eq. (7.16) at $t_0 = 0$ with using the Laplace transform and then multiplies the obtained expression by $\left(\vec{n}\left(\vec{r}_{S_{\beta^*}}\right)\cdot\vec{\Omega}\right)\left(\vec{m}\left(\vec{r}_{S_{\beta^*}}\right)\right)^{-2}$ and integrates over variables $\vec{r}$ and $\vec{\Omega}$ on $\mathrm{S}_{\beta^*}\left(\mathrm{V}^0_{\beta^*}\right)$ and on Ω, respectively, we get to the following relation

$$
\begin{aligned}
&\iint\limits_{\mathrm{S}_{\beta^*}\left(\mathrm{V}^0_{\beta^*}\right)} \left(m\left(\vec{r}\right)\right)^{-2} dS \int\limits_{\Omega} \left(\vec{n}\cdot\vec{\Omega}\right)\bar{\mathrm{I}}\left(\vec{r},\vec{\Omega};\sigma;\mathrm{V}_\beta\right) d\Omega \\
&= \iint\limits_{\mathrm{S}_{\beta^*}\left(\mathrm{V}^0_{\beta^*}\right)} \left(m\left(\vec{r}\right)\right)^{-2} dS \int\limits_{\Omega} \left(\vec{n}\cdot\vec{\Omega}\right) d\Omega \iiint\limits_{\mathrm{V}^0_\beta} d\mathrm{V}' \int\limits_{\Omega} \bar{\mathrm{G}}_*\left(\vec{r},\vec{\Omega};\vec{r}',\vec{\Omega}';\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right) \\
&\times\ \bar{g}^*_\beta\left(\vec{r}',\vec{\Omega}',\sigma;\mathrm{V}^0_\beta\right) d\Omega' - \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right)} dS' \int\limits_{\Omega} \left(\vec{n}'\cdot\vec{\Omega}'\right) \\
&\times \left[\iint\limits_{\mathrm{S}_{\beta^*}\left(\mathrm{V}^0_{\beta^*}\right)} \left(m\left(\vec{r}\right)\right)^{-2} dS \int\limits_{\Omega} \left(\vec{n}\cdot\vec{\Omega}\right)\bar{\mathrm{G}}_*\left(\vec{r},\vec{\Omega};\vec{r}',\vec{\Omega}';\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right) d\Omega\right] \\
&\times\ \bar{\mathrm{I}}\left(\vec{r}',\vec{\Omega}';\sigma;\mathrm{V}_\beta\right) d\Omega',
\end{aligned}
\tag{7.39}
$$

where $\bar{g}^*_\beta\left(\vec{r}',\vec{\Omega}',\sigma;\mathrm{V}^0_\beta\right) = \Theta_{\mathrm{V}^0_\beta}\left(\vec{r}\right)\left(\bar{g}_\beta\left(\vec{r},\vec{\Omega},\sigma;t;\mathrm{V}^0_\beta\right) + \left(v\left(\vec{r}\right)\right)^{-1}\mathrm{I}\left(\vec{r},\vec{\Omega};-0;\mathrm{V}_\beta\right)\right)$. The expression in the square brackets of Eq. (7.39) can be transformed to the below form with using the considerations similar to those utilized in subsection 7.3.3 and the accounting for the reciprocity principle (see Eq. (7.27)):

$$
\begin{aligned}
&\iint\limits_{\mathrm{S}_{\beta^*}\left(\mathrm{V}^0_{\beta^*}\right)} \left(m\left(\vec{r}\right)\right)^{-2} dS \int\limits_{\Omega} \left(\vec{n}\cdot\vec{\Omega}\right)\bar{\mathrm{G}}_*\left(\vec{r},\vec{\Omega};\vec{r}',\vec{\Omega}';\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right) d\Omega \\
&= -\left(m\left(\vec{r}'\right)\right)^{-4} \iiint\limits_{\mathrm{V}^0_{\beta^*}} \left(m\left(\vec{r}\right)\right)^2 d\mathrm{V} \int\limits_{\Omega} \bar{\mathrm{G}}_*\left(\vec{r}',-\vec{\Omega}';\vec{r},\vec{\Omega};\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right) \\
&\quad\times\ \left(\kappa^*_{ext}(\vec{r},\sigma) - \kappa_{sca}\left(\vec{r}\right)\right) d\Omega.
\end{aligned}
\tag{7.40}
$$

While deriving Eq. (7.40), one takes into account that radius-vector $\vec{r}'$ specifies a point outside of body V_{β^*} (more exactly, on $\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right)$). If one allows $\mathrm{S}_{\beta^*}\left(\mathrm{V}^0_{\beta^*}\right)$ to approach $\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right)$ by continuous deformation and takes into account the constancy of the refractive index on $\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right)$, then the below GIR for the Laplace image $\bar{\mathrm{F}}_1\left(\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right),\sigma\right)$ of the energy radiance $\mathrm{F}_1\left(\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right),t\right)$ follows from Eq. (7.39):

$$
\begin{aligned}
&\bar{\mathrm{F}}_1\left(\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right),\sigma\right) = \bar{\mathrm{F}}_2\left(\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right),\sigma;\ \mathrm{V}_{\beta_4}\right) \\
&+ \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right)} dS' \int\limits_{\Omega} \left(\vec{n}'\cdot\vec{\Omega}'\right)\bar{\mathrm{I}}^{\times\times}\left(\vec{r}',-\vec{\Omega}';\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right)\bar{\mathrm{I}}\left(\vec{r}',\vec{\Omega}';\sigma;\mathrm{V}_\beta\right) d\Omega',
\end{aligned}
\tag{7.41}
$$

where

$$\bar{\mathrm{F}}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right),\sigma;\mathrm{V}_{\beta_4}\right)$$
$$=\iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)} dS\int\limits_\Omega\left(\vec{n}\cdot\vec{\Omega}\right)d\Omega\iiint\limits_{\mathrm{V}_\beta^0} d\mathrm{V}'\int\limits_\Omega \bar{\mathrm{G}}_*\left(\vec{r},\vec{\Omega};\vec{r}\,',\vec{\Omega}';\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right)$$
$$\times\,\bar{g}_\beta^*\left(\vec{r}\,',\vec{\Omega}',\sigma;\mathrm{V}_\beta^0\right)d\Omega', \tag{7.42}$$

$$\bar{\mathrm{I}}^{\times\times}\left(\vec{r}\,',-\vec{\Omega}';\sigma;\mathrm{S}_{\beta_2},\mathrm{V}_{\beta_4}\right)$$
$$=\left(m\left(\vec{r}\,'\right)\right)^{-2}\iiint\limits_{\mathrm{V}_\beta^0}\left(m\left(\vec{r}\right)\right)^2\left(\kappa_{ext}^*(\vec{r};\sigma)-\kappa_{sca}\left(\vec{r}\right)\right)d\mathrm{V}$$
$$\times\int_\Omega \bar{\mathrm{G}}_*\left(\vec{r}\,',-\vec{\Omega}';\vec{r},\vec{\Omega};\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right)d\Omega. \tag{7.43}$$

At this, one should believe that $m\left(\vec{r}\,'\right)\equiv m=\text{const}$ on $\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)$ in Eq. (7.41). Function $\bar{\mathrm{I}}^{\times\times}\left(\vec{r}\,',-\vec{\Omega}';\sigma;\mathrm{S}_{\beta_2},\mathrm{V}_{\beta_4}\right)$ defined by Eq. (7.43) is a dimensionless quantity for any nonnegative σ, any $\vec{r}\,'$ and any $\vec{\Omega}'$. Besides, all its values belong to [0, 1]. Due to the importance of this conclusion, let its substantiation be given for deriving important consequences of the GIR of Eq. (7.41). First, function $\bar{\mathrm{G}}_*^1\left(\vec{r},\vec{\Omega};\vec{r}_1,\vec{\Omega}_1;\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4^*}\right)=m_1\left(\vec{r}_1,\vec{r}\right)\bar{\mathrm{G}}_*\left(\vec{r},\vec{\Omega};\vec{r}_1,\vec{\Omega}_1;\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right)$ is the Laplace image of the volume Green function for Eq. (7.1) for model medium $\mathrm{V}_{\beta_4^*}$. This medium differs from V_{β_4} only in that it has $m\left(\vec{r}\right)\equiv m=\text{const}$ for any $\vec{r}$ specifying points in $\mathrm{V}_{\beta_4^*}$ (but function $v\left(\vec{r}\right)$ should correspond to the real velocity of light in a turbid medium). The validity of the above interpretation of function $\bar{\mathrm{G}}_*^1\left(\vec{r},\vec{\Omega};\vec{r}_1,\vec{\Omega}_1;\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4^*}\right)$ directly follows from Eq. (7.1), when function $\bar{g}\left(\vec{r},\vec{\Omega},t;\mathrm{V}_\beta^0\right)$ is replaced in it by $\delta\left(\vec{r}-\vec{r}_1\right)\delta\left(\vec{\Omega}-\vec{\Omega}_1\right)\delta\left(t\right)$, and from the properties of the δ-functions. Note that the above interrelation between functions $\bar{\mathrm{G}}_*^1\left(\vec{r},\vec{\Omega};\vec{r}_1,\vec{\Omega}_1;\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4^*}\right)$ and $\bar{\mathrm{G}}_*\left(\vec{r},\vec{\Omega};\vec{r}_1,\vec{\Omega}_1;\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4^*}\right)$ is formally similar to Eq. (7.37). Second, the validity of the inequality below follows from this interrelation and Eq. (7.43), when σ is a nonnegative number:

$$\bar{\mathrm{I}}^{\times\times}\left(\vec{r}\,',-\vec{\Omega}';\sigma;\mathrm{S}_{\beta_2},\mathrm{V}_{\beta_4}\right)$$
$$\le\iiint\limits_{\mathrm{V}_\infty^*}\left(\kappa_{ext}^*(\vec{r},\sigma)-\kappa_{sca}\left(\vec{r}\right)\right)d\mathrm{V}\int_\Omega \bar{\mathrm{G}}_*^1\left(\vec{r}\,',-\vec{\Omega}';\vec{r},\vec{\Omega};\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right)d\Omega. \tag{7.44}$$

This inequality is valid for any $\vec{r}\,'$ specifying points in V_{β_4} and for any $\vec{\Omega}'\in\Omega$. In Eq. (7.44), function $\bar{\mathrm{G}}_*^1\left(\vec{r},\vec{\Omega};\vec{r}_1,\vec{\Omega}_1;\sigma;\mathrm{S}_{\beta_2};\mathrm{V}_\beta^*\right)$ should stand for the Laplace image of the Green function of Eq. (7.1) for the case of infinite model medium V_∞^*, in which $m\left(\vec{r}\right)\equiv m=\text{const}$ for any $\vec{r}$ specifying points in V_∞^*. Here, V_∞^* should comprise a portion that is identical to $\mathrm{V}_{\beta_4^*}$ by its optical and geometrical properties. Third, by using the considerations similar to those utilized in subsection 7.3.3, the expression in the right-hand side of Eq. (7.44) can be shown to be identically unity. Fourth, from the definition of function $\bar{\mathrm{I}}^{\times\times}\left(\vec{r}\,',-\vec{\Omega}';\sigma;\mathrm{S}_{\beta_2},\mathrm{V}_{\beta_4}\right)$ and from the

nonnegativity of σ, the nonnegativity of this function for any $\vec{r}'$ specifying points in V_{β_4} and for any $\vec{\Omega}' \in \Omega$ follows. The above sequence of considerations gives the validity of the two-sided inequality of $0 \le \bar{\mathrm{I}}^{\times\times}\left(\vec{r}', -\vec{\Omega}'; \sigma; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right) \le 1$ for any $\vec{r}'$ specifying points in V_{β_4} and for any $\vec{\Omega}'$ (but for nonnegative σ). Because V_β is not infinite turbid medium V_∞ according to the original assumption, then function $\bar{\mathrm{I}}^{\times\times}\left(\vec{r}', -\vec{\Omega}'; \sigma; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right)$ will assume values that are strictly less than unity for any above-said $\vec{r}'$ and $\vec{\Omega}'$.

The stationary analogs of Eqs (7.41)–(7.43) are the following formulae:

$$\mathrm{F}_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) = \mathrm{F}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right); \mathrm{V}_{\beta_4}\right) + \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)} dS' \int\limits_\Omega \left(\vec{n}' \cdot \vec{\Omega}'\right) \mathrm{I}^{\times\times}\left(\vec{r}', -\vec{\Omega}'; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right) \mathrm{I}\left(\vec{r}', \vec{\Omega}'; \mathrm{V}_\beta\right) d\Omega', \tag{7.45}$$

$$\mathrm{F}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0; \mathrm{V}_{\beta_4}\right)\right) = \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)} dS \int_\Omega \left(\vec{n} \cdot \vec{\Omega}\right) d\Omega \iiint\limits_{\mathrm{V}_\beta^0} d\mathrm{V}' \int\limits_\Omega \mathrm{G}_*\left(\vec{r}, \vec{\Omega}; \vec{r}', \vec{\Omega}'; \mathrm{S}_{\beta_2}; \mathrm{V}_{\beta_4}\right) \times g_\beta\left(\vec{r}', \vec{\Omega}'; \mathrm{V}_\beta^0\right) d\Omega', \tag{7.46}$$

$$\mathrm{I}^{\times\times}\left(\vec{r}', -\vec{\Omega}'; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right) = \iiint\limits_{\mathrm{V}_\beta^0} m_1\left(\overrightarrow{r}, \vec{r}'\right)\left(\kappa_{ext}(\vec{r}) - \kappa_{sca}\left(\vec{r}\right)\right) d\mathrm{V} \times \int\limits_\Omega \mathrm{G}_*\left(\vec{r}', -\vec{\Omega}'; \vec{r}, \vec{\Omega}; \mathrm{S}_{\beta_2}; \mathrm{V}_{\beta_4}\right) d\Omega, \tag{7.47}$$

where $m\left(\vec{r}'\right) \equiv m = \text{const}$ on $\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)$. Equation (7.45) is the GIR for monochromatic radiance $\mathrm{F}_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$. It equals the right-hand side of Eq. (7.38), when function $\mathrm{I}\left(\vec{r}, \vec{\Omega}; t, \mathrm{V}_\beta\right)$ is replaced by radiation intensity $\mathrm{I}\left(\vec{r}, \vec{\Omega}; \mathrm{V}_\beta\right)$ in body V_β that is illuminated by internal or external stationary radiation sources. The physical interpretation of the GIRs of Eqs (7.41) and (7.45), when $m\left(\vec{r}\right) \equiv m = \text{const}$ in V_β^0, was made by Rogovtsov (1986b).

7.3.5 Inequalities for monochromatic radiation fluxes and mean emission durations of turbid bodies

Note some consequences of the GIR of Eqs (7.41) and (7.45). Assume first that turbid medium V_β does not contain any underlying surfaces, has a non-concave shape, and is illuminated by internal stationary radiation sources only. Note that the investigation of the radiative transfer process in turbid media illuminated by external sources is reduced to studying of this process in the media comprising internal radiation sources of particular kinds (see, for example, the monographs (Sobolev, 1975; Yanovitskij, 1997)). Denote the least upper bound of the set of

values, which are assumed by function $\mathrm{I}^{\times\times}\left(\vec{r}', -\vec{\Omega}'; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right)$, with $\vec{r}'$ specifying points on $\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)$ and $\vec{\Omega}' \in \Omega_+\left(\vec{r}_{\mathrm{S}}'\right) = \Omega_+$, via $a(\mathrm{S}_\beta)$. Denote the greatest lower bound of the set of values, which are assumed by the same function with the same restraints of $\vec{r}'$ and $\vec{\Omega}'$, via $b(\mathrm{S}_\beta)$ $(b(\mathrm{S}_\beta) \le a(\mathrm{S}_\beta))$. Then, accounting for the absence of the external illumination and the definition of monochromatic radiance $\mathrm{F}_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$, one gets the following two-sided inequality from the GIR of Eq. (7.45):

$$\begin{aligned} \mathrm{F}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right); \mathrm{V}_{\beta_4}\right)\left(1 - b\left(\mathrm{S}_\beta\right)\right)^{-1} &\le \mathrm{F}_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \\ &\le \mathrm{F}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right); \mathrm{V}_{\beta_4}\right)\left(1 - a\left(\mathrm{S}_\beta\right)\right)^{-1} \end{aligned} \quad (7.48)$$

For turbid media, where there is no change in the refractive index (i. e. where $m(\vec{r}) \equiv m = \mathrm{const}$), various versions of the inequalities of Eq. (7.48) have been firstly derived by Rogovtsov (1983, 1985b, 1986b). It follows from the results of these papers that, for macroscopically homogeneous turbid media having constant or forward-extended phase functions, quantity $a(\mathrm{S}_\beta)$ does not exceed $\frac{1}{2}\omega_0$, where ω_0 is the single scattering albedo. At this, if the optical dimensions of homogeneous body V_β bounded by a smooth surface S_β with a non-zero curvature at any point do not change, then the quantities $a(\mathrm{S}_\beta)$ and $b(\mathrm{S}_\beta)$ will decrease as the forward extension of the phase function increases. It follows from the above conclusions and from Eq. (7.48) that, for macroscopically homogeneous turbid media having constant or forward-extended phase functions, the simple inequalities below are valid (Rogovtsov, 1986b):

$$\mathrm{F}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right); \mathrm{V}_\infty\right) \le \mathrm{F}_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \le \left(1 - 2^{-1}\omega_0\right)^{-1} \mathrm{F}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right); \mathrm{V}_\infty\right). \quad (7.49)$$

A macroscopically homogeneous infinite medium having the same optical characteristics as those of body V_β should stand for V_∞ of Eq. (7.49). It follows immediately from Eq. (7.49) that quantity $\mathrm{F}_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$ differs from $\mathrm{F}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right); \mathrm{V}_\infty\right)$ by not more than two times for any albedo values ω_0 $(\omega_0 \in [0, 1])$. Monochromatic flux $\mathrm{F}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right); \mathrm{V}_\infty\right)$ can be generally found for any phase functions and any isotropic internal sources in body V_β bounded by any non-concave surface S_β. This can be made by, for example, the results of the papers (Rogovtsov, 1996, 1997; Rogovtsov and Borovik, 2009).

Now let a non-concave body V_β comprises no underlying surfaces and contains a point mono-directional non-stationary radiation source with density $\varphi(t)\,\delta(\vec{r} - \vec{r}_1)$ $\delta\left(\vec{\Omega} - \vec{\Omega}_1\right)$. We will assume that there was no radiation field in body V_β before time moment $t = 0$ and that function $\varphi(t)$ satisfies the following conditions, namely (i) $\varphi(t)$ is a piecewise smooth function; (ii) $\varphi'(t) \ge 0$ for any $t \ge 0$; and (iii) there exists the limit of $\lim_{t\to+\infty} \varphi(t) = \varphi_0$ $(\varphi_0 = \mathrm{const} > 0)$. Under the above assumptions, one can get the following inequalities from the GIR of Eq. (7.41) and from the inequalities of Eqs (7.48) with using the Laplace transform:

$$\begin{aligned} \mathrm{F}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right); t, \mathrm{V}_{\beta_4}\right) &\le \mathrm{F}_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right), t\right) \\ &\le \mathrm{F}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right); +\infty; \mathrm{V}_{\beta_4}\right)\left(1 - a\left(\mathrm{S}_\beta\right)\right)^{-1}. \end{aligned} \quad (7.50)$$

The inequalities of Eq. (7.50) are valid for any $t > 0$. A monochromatic flux through surface S_β, when there is a stationary point mono-directional radiation source with

density $\varphi_0\delta\left(\vec{r}-\vec{r}_1\right)\delta\left(\vec{\Omega}-\vec{\Omega}_1\right)$ in portion V_β of body V_{β_4}, should be understood as $\mathrm{F}_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_{\beta_4}\right) = \varphi_0 \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)} dS \int_\Omega \left(\vec{n}\cdot\vec{\Omega}\right) G_*\left(\vec{r},\vec{\Omega};\vec{r}_1,\vec{\Omega}_1;\mathrm{S}_{\beta_2};\mathrm{V}_{\beta_4}\right) d\Omega$ where $\mathrm{S}_{\beta_2} = \mathrm{S}_\beta$. The inequalities of Eq. (7.50) enable, in principle, the monochromatic radiance $\mathrm{F}_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right),t\right)$ of body V_β to be analytically estimated, when V_β can be embedded in an infinite macroscopically homogeneous medium V_∞.

The derivation of exact or approximate analytical solutions to the boundary value problems of Eq. (7.1) for media with rather complicated configurations is quite a labor-consuming problem. Therefore, it is sometimes useful to find various averaged (integral) characteristics of radiation fields in such media. One can take, as such characteristics, for example, the time moments of monochromatic radiance $\mathrm{F}_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right),t\right)$ of body V_β, through which the mean emission durations of body V_β are expressed under the illumination by non-stationary radiation sources or under the interruption of such an illumination at some instant of time. The papers of Irvine (1968), Katsev (1969), Rogovtsov and Samson (1975) are devoted to the design of methods for calculating the time moments for various non-stationary radiation fields in macroscopically homogeneous absorbing and scattering media. The general expressions and inequalities written down in subsections 7.3.3 and 7.3.4 as well as in this subsection enable one to get a number of strict and asymptotic formulae and inequalities for the above-said time moments and mean emission durations t^*.

Let all the conditions imposed on the properties of bodies V_β and V_{β_4}, while deriving the GIR of Eq. (7.41) be satisfied. Besides, we will assume that the zero and first time moments of functions, which have been obtained by the integration of the densities of internal and external sources over spatial (within V_β) and angular (within Ω), are finite quantities. At this, the below integrals will be stood for time moments $M_r\left(f\right)$ of order r from function $f(t)$defined at $t>0$

$$M_r\left(f\right) = \int_{-\varepsilon}^{+\infty} t^r f\left(t\right)\, dt;\ r=0,1,2,\ldots. \tag{7.51}$$

The quantity below will be understood as the mean emission duration

$$t^* = M_1(f)\left(M_0(f)\right)^{-1} \quad . \tag{7.52}$$

If one takes the monochromatic radiance $\mathrm{F}_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right),t\right)$ of body V_β as function $f(t)$, then we can get the following relations using the GIR of Eq. (7.41), the Laplace transform, and Eqs (7.40)–(7.43), (7.47), (7.51), and (7.1)

$$\begin{aligned} M_0(F_1) \;=\; & U_0 + \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)} dS' \int_\Omega \left(\vec{n}'\cdot\vec{\Omega}'\right) \mathrm{I}^{\times\times}\left(\vec{r}',-\vec{\Omega}';\mathrm{S}_{\beta_2},\mathrm{V}_{\beta_4}\right) d\Omega' \\ & \times \int_{-\varepsilon}^{+\infty} \mathrm{I}\left(\vec{r}',\vec{\Omega}',t;\mathrm{V}_\beta\right) dt, \end{aligned} \tag{7.53}$$

$$M_1(F_1) \;=\; U_1 + \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)} dS' \int\limits_{\Omega} \left(\vec{n}'\cdot\vec{\Omega}'\right)\mathrm{I}^{\times\times}\left(\vec{r}', -\vec{\Omega}'; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right) d\Omega' \times \int\limits_0^{+\infty} \mathrm{tI}\left(\vec{r}', \vec{\Omega}', t; \mathrm{V}_\beta\right) dt, \tag{7.54}$$

where

$$U_0 \;=\; \iiint\limits_{\mathrm{V}_\beta^0} m_1\left(\vec{r}_{\mathrm{S}_\beta}, \vec{r}'\right) d\mathrm{V}' \int\limits_{\Omega} \chi_0\left(\vec{r}', \vec{\Omega}'; \mathrm{V}_\beta^0\right)\left(1 - \mathrm{I}^{\times\times}\left(\vec{r}', -\vec{\Omega}'; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right)\right) d\Omega', \tag{7.55}$$

$$\begin{aligned} U_1 \;=\; & \iiint\limits_{\mathrm{V}_\beta^0} m_1\left(\vec{r}_{\mathrm{S}_\beta}, \vec{r}'\right) d\mathrm{V}' \int\limits_{\Omega} \chi_1\left(\vec{r}', \vec{\Omega}'; \mathrm{V}_\beta^0\right) \Lambda\left(\vec{r}', -\vec{\Omega}'; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right) d\Omega' \\ & + \iiint\limits_{\mathrm{V}_\beta^0} m_1\left(\vec{r}_{\mathrm{S}_\beta}, \vec{r}'\right) d\mathrm{V}' \int\limits_{\Omega} \chi_0\left(\vec{r}', \vec{\Omega}'; \mathrm{V}_\beta^0\right) d\Omega' \iiint\limits_{\mathrm{V}_{\beta_4}} m_1\left(\vec{r}'', \vec{r}'\right) \frac{d\mathrm{V}''}{v\left(\vec{r}''\right)} \\ & \times \int\limits_{\Omega} \mathrm{G}_*\left(\vec{r}', -\vec{\Omega}'; \vec{r}'', \vec{\Omega}''; \mathrm{S}_{\beta_2}; \mathrm{V}_{\beta_4}\right) \Lambda\left(\vec{r}'', \vec{\Omega}''; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right) d\Omega'' \\ & - \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)} dS' \int\limits_{\Omega} \left(\vec{n}'\cdot\vec{\Omega}'\right) d\Omega' \int\limits_{-\varepsilon}^{+\infty} \mathrm{I}\left(\vec{r}', \vec{\Omega}', t; \mathrm{V}_\beta\right) dt \iiint\limits_{\mathrm{V}_{\beta_4}} m_1\left(\vec{r}'', \vec{r}'\right) \frac{d\mathrm{V}''}{v\left(\vec{r}''\right)} \\ & \times \int\limits_{\Omega} \mathrm{G}_*\left(\vec{r}', -\vec{\Omega}'; \vec{r}'', \vec{\Omega}''; \mathrm{S}_{\beta_2}; \mathrm{V}_{\beta_4}\right) \Lambda\left(\vec{r}'', \vec{\Omega}''; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right) d\Omega''. \end{aligned} \tag{7.56}$$

Functions $\chi_0\left(\vec{r}', \vec{\Omega}'; \mathrm{V}_\beta^0\right)$, $\chi_1\left(\vec{r}', \vec{\Omega}'; \mathrm{V}_\beta^0\right)$, $\Lambda\left(\vec{r}', \vec{\Omega}'; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right)$ entering Eqs (7.55) and (7.56) are defined by the following expressions:

$$\begin{aligned} & \chi_0\left(\vec{r}', \vec{\Omega}'; \mathrm{V}_\beta^0\right) = \Theta_{\mathrm{V}_\beta^0}\left(\vec{r}'\right)\left[\int\limits_{-\varepsilon}^{+\infty} g_\beta\left(\vec{r}', \vec{\Omega}', t; \mathrm{V}_\beta\right) dt + \left(v\left(\vec{r}'\right)\right)^{-1} I\left(\vec{r}', \vec{\Omega}', -0; \mathrm{V}_\beta\right)\right], \\ & \chi_1\left(\vec{r}', \vec{\Omega}'; \mathrm{V}_\beta^0\right) = \Theta_{\mathrm{V}_\beta^0}\left(\vec{r}'\right) \int\limits_0^{+\infty} t g_\beta\left(\vec{r}', \vec{\Omega}', t; \mathrm{V}_\beta^0\right) dt, \\ & \Lambda\left(\vec{r}', \vec{\Omega}'; \mathrm{S}_{\beta_2}, \mathrm{V}_{\beta_4}\right) = \Theta_{\mathrm{V}_\beta^0}\left(\vec{r}'\right) - \mathrm{I}^{\times\times}\left(\vec{r}', \vec{\Omega}'; \mathrm{S}_{\beta_2}; \mathrm{V}_{\beta_4}\right). \end{aligned} \tag{7.57}$$

A number of strict expressions, inequalities, or asymptotic formulae for mean emission duration t^* of turbid bodies of various configurations can be derived by using Eqs (7.52)–(1.57). Below are a number of such examples.

Let body V_β have no underlying surfaces and be a non-concave (recall that the identity of $m\left(\vec{r}_{\mathrm{S}_\beta}\right) \equiv m = \mathrm{const}$ is valid on the boundary S_β of body V_β) and a conservatively scattering one (i.e. $\kappa_{ext}\left(\vec{r}\right) \equiv \kappa_{sca}\left(\vec{r}\right)$). Besides, assume that,

starting from time moment $t = -\varepsilon$, the turbid medium is illuminated by internal radiation sources only or is not illuminated at all. Then one gets the following relation from Eqs (7.52) and (7.53), (7.54), when the body V_β is taken as V_{β_4}:

$$
\begin{aligned}
t^* \;=\; & \left[\iiint\limits_{\mathrm{V}_\beta^0} m_1\left(\vec{r}_{\mathrm{S}_\beta}, \vec{r}\,'\right) d\mathrm{V}' \int\limits_{\Omega} \chi_1\left(\vec{r}\,', \vec{\Omega}'; \mathrm{V}_\beta^0\right) d\Omega' \right. \\
& + \iiint\limits_{\mathrm{V}_\beta^0} m_1\left(\vec{r}_{\mathrm{S}_\beta}, \vec{r}\,'\right) d\mathrm{V}' \int\limits_{\Omega} \chi_0\left(\vec{r}\,', \vec{\Omega}'; \mathrm{V}_\beta^0\right) d\Omega' \\
& \left. \times \iiint\limits_{\mathrm{V}_\beta} m_1\left(\vec{r}\,'', \vec{r}\,'\right) \frac{d\mathrm{V}''}{v\left(\vec{r}\,''\right)} \int\limits_{\Omega} \mathrm{G}_*\left(\vec{r}\,', -\vec{\Omega}'; \vec{r}\,'', \vec{\Omega}''; \mathrm{S}_\beta; \mathrm{V}_\beta\right) d\Omega'' \right] \\
& \times \left[\iiint\limits_{\mathrm{V}_\beta^0} m_1\left(\vec{r}_{\mathrm{S}_\beta}, \vec{r}\,'\right) d\mathrm{V}' \int\limits_{\Omega} \chi_0\left(\vec{r}\,', \vec{\Omega}'; \mathrm{V}_\beta^0\right) d\Omega' \right]^{-1}. \qquad (7.58)
\end{aligned}
$$

It follows from Eq. (7.58) that, for the above conditions, the volume Green function of Eq. (7.10), which corresponds to the case of stationary sources, is enough to be found for calculating t^*. A particular case of Eq. (7.58), when $m\left(\vec{r}\right) \equiv m =$ const in the whole body V_o, has been obtained in the papers (Rogovtsov, 1985c; Rogovtsov and Samson, 1985a). Eq. (7.58) can be, in particular, utilized for deriving the asymptotic relations for quantity t^*, when turbid medium V_β is an optically thick one. Optically thick turbid medium V_β everywhere below stands for such a medium that comprises regions positioned at large optical distances from all points of boundary S_β of body V_β. In its own turn, optical distance $\tau_0\left(\vec{r}_0, \vec{r}_{\mathrm{S}_\beta}\right)$ from a point specified by $\vec{r}_0$ to a point lying on S_β and specified by $\vec{r}_{\mathrm{S}_\beta}$ is stood for integral $\int_0^{\left|\vec{r}_0 - \vec{r}_{\mathrm{S}_\beta}\right|} \kappa_{ext}\left(\vec{r}_0 - \xi(\vec{r}_0 - \vec{r}_{\mathrm{S}_\beta}) \left|\vec{r}_0 - \vec{r}_{\mathrm{S}_\beta}\right|^{-1}\right) d\xi$. A number of certain examples of using Eq. (7.58) have been given in the papers (Rogovtsov, 1985c; Rogovtsov and Samson, 1985a).

Let body V_β satisfy to the conditions of the preceding paragraph and have a finite optical thickness along any straight line intersecting it. Take a totally radiation-transparent surface as S_{β_2} and infinite turbid medium V_∞ as body V_{β_4}. Within the frame of these assumptions, it follows from Eqs (7.52)–(7.57) the validity of the below formula for the mean emission time of body V_β:

$$
t^* = \left(U_0\right)^{-1} U_1 \qquad (7.59)
$$

$$
U_0 = \iiint\limits_{\mathrm{V}_\beta^0} m_1\left(\vec{r}_{\mathrm{S}_\beta}, \vec{r}\,'\right) d\mathrm{V}' \int\limits_{\Omega} \chi_0\left(\vec{r}\,', \vec{\Omega}'; \mathrm{V}_\beta^0\right) d\Omega', \qquad (7.60)
$$

$$
\begin{aligned}
U_1 \;=\; & \iiint\limits_{\mathrm{V}_\beta^0} m_1\left(\vec{r}_{\mathrm{S}_\beta},\vec{r}\,'\right)d\mathrm{V}' \int\limits_{\Omega} \chi_1\left(\vec{r}\,',\vec{\Omega}';\mathrm{V}_\beta^0\right)d\Omega' \\
& + \iiint\limits_{\mathrm{V}_\beta^0} m_1\left(\vec{r}_{\mathrm{S}_\beta},\vec{r}\,'\right)\,d\mathrm{V}' \int\limits_{\Omega} \chi_0\left(\vec{r}\,',\vec{\Omega}';\mathrm{V}_\beta^0\right)d\Omega' \\
& \times \iiint\limits_{\mathrm{V}_\beta} m_1\left(\vec{r}\,'',\vec{r}\,'\right)\frac{d\mathrm{V}''}{v\left(\vec{r}\,''\right)} \int\limits_{\Omega} \mathrm{G}_*\left(\vec{r}\,',-\vec{\Omega}';\vec{r}\,'',\vec{\Omega}'';\mathrm{S}_{\beta_2};\mathrm{V}_\infty\right)d\Omega'' \\
& - \iint\limits_{\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)} dS' \int\limits_{\Omega} \left(\vec{n}'\cdot\vec{\Omega}'\right)\bar{\mathrm{I}}\left(\vec{r}\,',\vec{\Omega}',0;\mathrm{V}_\beta\right)d\Omega' \\
& \times \iiint\limits_{\mathrm{V}_\beta^0} m_1\left(\vec{r}\,'',\vec{r}\,'\right)\frac{d\mathrm{V}''}{v\left(\vec{r}\,''\right)} \int\limits_{\Omega} \mathrm{G}_*\left(\vec{r}\,',-\vec{\Omega}';\vec{r}\,'',\vec{\Omega}'';\mathrm{S}_{\beta_2};\mathrm{V}_\infty\right)d\Omega''. \qquad (7.61)
\end{aligned}
$$

In Eq. (7.61), function $\bar{\mathrm{I}}\left(\vec{r}\,',\vec{\Omega}',0;\mathrm{V}_\beta\right)=\int_{-\varepsilon}^{+\infty}\mathrm{I}\left(\vec{r}\,'_{\mathrm{S}_\beta},\vec{\Omega}',t;\mathrm{V}_\beta\right)dt$ with $\vec{\Omega}\in\Omega_+\left(\vec{r}_{\mathrm{S}_\beta}\right)$ is the reduction of the solution to Eq. (7.5) at $\sigma=0$ to boundary S_β of body V_β. At this, the said equation should be solved taking account of the boundary condition of $\bar{\mathrm{I}}\left(\vec{r}\,',\vec{\Omega}',0;\mathrm{V}_\beta\right)=0$ for any $\vec{\Omega}\in\Omega_-\left(\vec{r}_{\mathrm{S}_\beta}\right)$ and any $\vec{r}_{\mathrm{S}_\beta}$ specifying points on S_β. In fact, Eqs (7.59)–(7.61) reduce the problem of calculating the mean emission duration t^* of a turbid, conservatively scattering medium with any non-concave shape and with a finite optical thickness to the derivation of two functions. The first one is the volume Green function for RTE of Eq. (7.10) for an infinite medium V_∞ comprising a point isotropic radiation source. The second function is the radiation intensity exiting body V_β. At this, the body comprises internal sources that are specified by density $\alpha\left[\bar{g}\left(\vec{r},\vec{\Omega},0;\mathrm{V}_\beta^0\right)+\left(v\left(\vec{r}\right)\right)^{-1}I\left(\vec{r},\vec{\Omega},-0;\mathrm{V}_\beta\right)\right]\Theta_{\mathrm{V}_\beta^0}\left(\vec{r}\right)$, where α is a quantity being equal to unity and having the dimension reciprocal to time. Note that analytical expressions have been derived for the said Green function for the case of an infinite macroscopically homogeneous turbid medium in the literature (Kolesov, 1983; Rogovtsov, 1986a, 1986b). It should also be emphasized that to find functions $\mathrm{G}_*\left(\vec{r},\vec{\Omega};\vec{r}_1,\vec{\Omega}_1;\mathrm{S}_{\beta_2};\mathrm{V}_\infty\right)$ and $\bar{\mathrm{I}}\left(\vec{r}_{\mathrm{S}_\beta},\vec{\Omega},0;\mathrm{V}_\beta\right)$ entering Eqs (7.59) and (7.61) is a substantially simpler problem than to find time moment $M_1(F_1)$ and quantity t^* on the basis of the direct solution to the original boundary value problem for Eq. (7.1).

Consider now a non-concave, non-conservatively scattering turbid medium V_β having no underlying surfaces and not being illuminated by external radiation sources at $t>-\varepsilon$. There will be imposed no restraints on the optical dimensions of body V_β. Under these conditions, one gets the following two-sided inequalities from Eqs (7.52)–(7.54):

$$
\left(1-a\left(\mathrm{S}_\beta\right)\right)U_1\left(\left(1-b\left(\mathrm{S}_\beta\right)\right)U_0\right)^{-1}\le t^*\le\left(1-b\left(\mathrm{S}_\beta\right)\right)U_1\left(\left(1-a\left(\mathrm{S}_\beta\right)\right)U_0\right)^{-1}. \qquad (7.62)
$$

If $m\left(\vec{r}\right)\equiv m=\mathrm{const}$ at any point of turbid medium V_β, then the inequalities of Eq. (7.62) come to the inequalities derived earlier by Rogovtsov (1983, 1994). In these works, there were derived a number of useful consequences of the particular

cases of the GIRs of Eqs (7.41) and (7.45) and of Eqs (7.48), (7.49), and (7.62) as applied to macroscopically homogeneous turbid media of various configurations.

Two-sided inequalities have only been written down for monochromatic fluxes (radiances) and mean emission durations of turbid bodies. One can derive the two-sided inequalities for radiation intensities, mean radiation intensities, mean numbers of photon scattering events, and for a number of other quantities by using the GIRs and the above inequalities. Additional evidence referring to the derivation of the two-sided inequalities for these quantities are given, in particular, in Rogovtsov (1985b, 2002) and Rogovtsov and Samson (1987).

7.4 Strict, asymptotic and approximate analytical solutions to boundary-value problems of the radiative transfer theory for turbid media of various configurations

The derivation of exact, asymptotic and approximate analytical solutions (but rather correct solutions having an *a priori* estimation of correctness) to boundary-value problems of the radiative transfer theory for the case of arbitrary phase functions and turbid media of various configurations is connected with finding the solutions to rather comprehensive general and particular mathematical problems. The general invariance relations reduction method enables one to cope with these problems due to its commonality and heuristicity. A whole number of examples of such a kind will be given in this section. At this, we will mainly focus our attention at the establishment of final results obtained on the basis of using the GIRRM for solving specific boundary-value problems of the RTT for the case of macroscopically homogeneous turbid media. Besides, the refractive index $m\,(\vec{r})$ will be believed everywhere below to be independent of an 'observation' point in turbid media (i.e. $m\,(\vec{r}) \equiv m = \text{const}$).

7.4.1 Application of the general invariance relations reduction method to the derivation of azimuth-averaged reflection function for a macroscopically homogeneous plane-parallel semi-infinite turbid medium

Let a semi-infinite macroscopically homogeneous isotropic turbid medium $\mathrm{V}_{[0,+\infty)}$ be illuminated by a mono-directional, infinitely wide beam of external radiation. Assume that the density of the incident light flux on the area perpendicular to the beam is equal to F ($F \equiv \text{const}$) at the top of the turbid medium. Hence, the intensity of the external radiation is equal to $F\delta\big(\vec{\Omega} - \vec{\Omega}_1\big)$, where unit vectors $\vec{\Omega}$ and $\vec{\Omega}_1$ specify the radiation propagation direction in medium $\mathrm{V}_{[0,+\infty)}$ and external radiation beam, respectively. Radiation intensity $\mathrm{I}\big(\vec{r},\vec{\Omega};\mathrm{V}_{[0,+\infty)}\big)$ in turbid medium $\mathrm{V}_{[0,+\infty)}\big(\mathrm{I}\big(\vec{r},\vec{\Omega};\mathrm{V}_{[0,+\infty)}\big) \equiv \mathrm{I}\big(z,\vec{\Omega};\vec{\Omega}_1;\mathrm{V}_{[0,+\infty)}\big)\big)$ is the solution to the following boundary-value problem:

$$\mu\frac{\partial I\left(z,\vec{\Omega};\vec{\Omega}_1;\mathrm{V}_{[0,+\infty)}\right)}{\partial \mathrm{z}} = -\kappa_{ext}I\left(\vec{z},\vec{\Omega};\vec{\Omega}_1;\mathrm{V}_{[0,+\infty)}\right) + \frac{\kappa_{sca}}{4\pi}\int\limits_{\Omega} p\left(\vec{\Omega}\cdot\vec{\Omega}'\right)I\left(\vec{z},\vec{\Omega}';\vec{\Omega}_1;\mathrm{V}_{[0,+\infty)}\right)d\Omega'; \quad (7.63)$$

$$I\left(+0,\vec{\Omega};\vec{\Omega}_1;\mathrm{V}_{[0,+\infty)}\right)\Big|_{\mu_1\in(0,1]} = F\delta\left(\vec{\Omega}-\vec{\Omega}_1\right). \quad (7.64)$$

Here $z \in (0,+\infty)$; $\kappa_{ext}\left(\vec{r}\right) \equiv \kappa_{ext} = \text{const}$; $\kappa_{sca}\left(\vec{r}\right) \equiv \kappa_{sca} = \text{const}$; $p\left(\vec{r};\vec{\Omega}\cdot\vec{\Omega}'\right) \equiv p\left(\vec{\Omega}\cdot\vec{\Omega}'\right)$; z is the depth of an 'observation' point inside medium $\mathrm{V}_{[0,+\infty)}$ (z is the shortest distance from the 'observation' point to boundary $S_{[0,+\infty)}$ of medium $\mathrm{V}_{[0,+\infty)}$; $\mu_1 = -\left(\vec{n}\cdot\vec{\Omega}_1\right)$, where $\vec{n}$ is the unit external normal to $S_{[0,+\infty)}$. The radiation intensity exiting through boundary $S_{[0,+\infty)}$ can be expressed via the reflection function $R\left(\vec{\Omega},\vec{\Omega}_1\right)$ (Sobolev, 1975; Yanovitskij, 1997) by means of the relation:

$$I\left(0,\vec{\Omega};\vec{\Omega}_1;\mathrm{V}_{[0,+\infty)}\right)\Big|_{\mu_1\in[-1,0)} = \frac{F}{\pi}R\left(\vec{\Omega},\vec{\Omega}_1\right)\mu_1, \quad (7.65)$$

where $\mu = -\left(\vec{n}\cdot\vec{\Omega}\right)$, $R\left(\vec{\Omega},\vec{\Omega}_1\right) = R\left(|\mu|,\varphi,\mu_1,\varphi_1\right)$ (φ and φ_1 are the azimuths corresponding to vectors $\vec{\Omega}$ and $\vec{\Omega}_1$). The physical sense of the reflection function $R\left(|\mu|,\varphi,\mu_1,\varphi_1\right)$ is explained in detail by Sobolev (1975) and Kokhanovsky (2006).

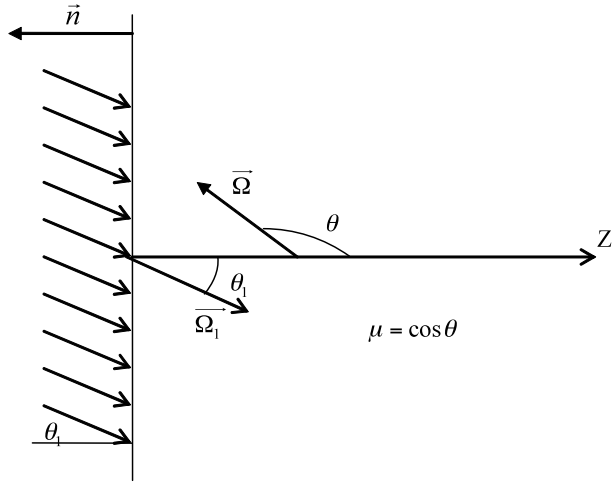


Fig. 7.6. Geometry of problem

Fig. 7.6 gives the graphical explanations of the above-introduced quantities. Quantity $I^0\left(z,\mu;\mu_1;\mathrm{V}_{[0,+\infty)}\right) = (2\pi)^{-1}\int_0^{2\pi} I\left(z,\vec{\Omega};\vec{\Omega}_1;\mathrm{V}_{[0,+\infty)}\right)d\varphi$ that due to the isotropic local characteristics of medium $\mathrm{V}_{[0,+\infty)}$ is independent of azimuth φ_1 will stand for the azimuth-averaged radiation intensity. Azimuth-averaged radiation intensity $I^0\left(0,-|\mu|;\mu_1;\mathrm{V}_{[0,+\infty)}\right)$ exiting $\mathrm{V}_{[0,+\infty)}$ is expressed via azimuth-averaged reflection function $R^0\left(|\mu|,\mu_1\right) = (2\pi)^{-1}\int_0^{2\pi} R\left(|\mu|,\varphi,\mu_1,\varphi_1\right)d\varphi$ by the relation of

$I^0\left(0,-|\mu|;\mu_1;\mathrm{V}_{[0,+\infty)}\right)=\frac{F}{\pi}R^0\left(|\mu|,\mu_1\right)\mu_1$. Function $R^0\left(|\mu|,\mu_1\right)$ can be found in essentially analytical form for any albedo $\omega_0\left(\omega_0=\kappa_{sca}\left(\kappa_{ext}\right)^{-1}\right)$ and any phase function $p\left(\mu\right)$ $\left(\mu\in[-1,1]\right)$ by using the particular version of the GIR of Eq. (7.21). The general expression for this function for the case of the conservative scattering $(\omega_0=1)$ and the calculation algorithm have been obtained and given in detail by Rogovtsov and Borovik (2009).

To describe briefly the simplest version of the method for deriving $R^0\left(|\mu|,\mu_1\right)$ for the case of non-conservative scattering, we will start from the particular version of the GIR of Eq. (7.21) that has the following form (Rogovtsov and Borovik, 2009):

$$\begin{aligned}R^0\left(|\mu|,\mu_1,\omega_0\right) &= \frac{1}{2}\tilde{G}_{\infty;0}\left(+0,-|\mu|;0,\mu_1;\omega_0\right)\\ &\quad-\int\limits_0^1\mu'\tilde{G}_{\infty;0}\left(+0,-|\mu|;0,-\mu';\omega_0\right)R^0\left(\mu',\mu_1,\omega_0\right)\,d\mu'. \quad (7.66)\end{aligned}$$

Here $|\mu|\in(0,1]$, $\mu_1\in(0,1]$; $R^0\left(|\mu|,\mu_1,\omega_0\right)$ is the azimuth-averaged reflection function corresponding to single scattering albedo ω_0 $(\omega_0\in(0,1))$; $\tilde{G}_{\infty;0}\left(\tau,\mu;\tau_1,\mu_1;\omega_0\right)$ is the Green function of the dimensionless RTE for the case of an macroscopically homogeneous isotropic infinite turbid medium V_∞ comprising a 'source' with density $\delta\left(\tau-\tau_1\right)\delta\left(\mu-\mu_1\right)$ (τ and τ_1 are the optical thickness values, where the 'observation' point and the 'source' are located, respectively). The said Green function is the solution to the following problem:

$$\begin{aligned}\mu\frac{\partial\tilde{G}_{\infty;0}\left(\tau,\mu;\tau_1,\mu_1;\omega_0\right)}{\partial\tau} &= -\tilde{G}_{\infty;0}\left(\tau,\mu;\tau_1,\mu_1;\omega_0\right)\\ &\quad+\frac{\omega_0}{2}\int\limits_{-1}^1 P_0\left(\mu,\mu'\right)\tilde{G}_{\infty;0}\left(\tau,\mu';\tau_1,\mu_1;\omega_0\right)\,d\mu'\\ &\quad+\delta\left(\tau-\tau_1\right)\delta\left(\mu-\mu_1\right). \quad (7.67)\end{aligned}$$

$$\begin{aligned}&\lim_{|\tau-\tau_1|\to+\infty}\tilde{G}_{\infty;0}\left(\tau,\mu;\tau_1,\mu_1;\omega_0\right)=0,\ \tau,\tau_1\in(-\infty,+\infty),\ \mu\in[-1,1],\\ &\mu_1\in[-1,0)\cup(0,1]. \quad (7.68)\end{aligned}$$

Function $P_0\left(\mu,\mu'\right)$ of Eq. (7.27) is defined by the relation

$$P_0\left(\mu,\mu'\right)=\sum_{S=0}^{+\infty}\left(2s+1\right)f_sP_s\left(\mu\right)P_s\left(\mu'\right),\ \left(\mu,\mu'\right)\in[-1,1]\times[-1,1], \quad (7.69)$$

where $(2s+1)\,f_s$ are the coefficients of the series expansion of phase function $p\left(\mu\right)$ in Legendre polynomials $P_s\left(\mu\right)$, i.e. $p\left(\mu\right)=\sum_{s=0}^{+\infty}\left(2s+1\right)f_sP_s\left(\mu\right)$. While getting Eq. (7.66), the local optical characteristics of media $\mathrm{V}_{[0,+\infty)}$ and V_∞ are assumed to coincide with each other. The relations of the kind of Eq. (7.66) have been first derived by independent and different ways by Ivanov (1976), Rogovtsov and Samson (1976) and Domke (1976). In this respect, Rogovtsov and Samson (1976) and Domke (1976) have shown that invariance relations of the kind of Eq. (7.66) can

be effectively utilized to find the radiation intensity exiting plane-parallel media. The detailed physical interpretation of these relations has been given by Ivanov and Volkov (1978). The invariance relations of Eq. (7.66) is an integral equation (IE) with respect to unknown function $R^0(|\mu|, \mu_1, \omega_0)$, if one assumes that the Green function $\tilde{G}_{\infty;0}(\tau, \mu; 0, \mu_1; \omega_0)$ is known, when $\tau = +0$, $\mu = -|\mu|$ and $\mu_1 \in [-1, 0) \cup (0, 1]$. The analytical expressions for this Green function for the case of an arbitrary phase function have been derived by Rogovtsov and Borovik (1993, 2009). Furthermore, in Rogovtsov (1996) and Rogovtsov and Borovik (2009) the algorithms for calculating the said Green function, the discrete roots, and the eigenfunctions of the reduced characteristic equations of the radiative transfer theory have been proposed.

Now use the results of the papers (Rogovtsov, 1996; Rogovtsov and Borovik, 2009) and the reduction procedure to reduce the IE of Eq. (7.66) to another integral equation that can be effectively solved by the iteration method. To do this, transform the integral equation of Eq. (7.66) taking account of the representation below of the Green function $\tilde{G}_{\infty;0}(\tau, \mu; 0, \mu'; \omega_0)$ (Rogovtsov and Borovik, 2009):

$$\begin{aligned}\tilde{G}_{\infty;0}(\tau, \mu; 0, \mu'; \omega_0) &= c_1(\omega_0; 0)\, \Phi_0(-ik_1(0), \mu)\, \Phi_0(-ik_1(0), \mu') \exp(-k_1(0)\tau) \\ &\quad + \tilde{G}^*_{\infty;0}(\tau, \mu; 0, \mu'; \omega_0), \quad \tau > 0. \end{aligned} \tag{7.70}$$

Here $\Phi_0(-ik_1(0), \mu)$ is the eigenfunction of the reduced characteristic equations of the RTT corresponding to the zeroth azimuthal harmonics (see, for example, Rogovtsov, 1996); $k_1(0)$ is the least positive root of the equation of $\wp_0(-k^2; 0) = 0$, where function

$$\wp_0(-k^2; 0) = \left[1; \frac{-q_0(0)\, k^2}{1}; \frac{-q_1(0)\, k^2}{1}, \ldots\right]$$

is the infinite continued fraction (see, for example, Jones and Thron (1980) and Rogovtsov and Borovik (2009)), in which $q_l(0) = (l+1)^2 \left[(2l+1)(2l+3)(1-\omega_0 f_l)(1-\omega_0 f_{l+1})\right]^{-1}$ for all $l = 0, 1, 2, \ldots$; $c_1(\omega_0; 0) = \left[\int_{-1}^{1} \mu \left(\Phi_0(-ik_1(0), \mu)\right)^2 d\mu\right]^{-1}$ is the normalization constant (here, the function itself $\Phi_0(-ik_1(0), \mu)$ is normalized by the condition of $\int_{-1}^{1} \Phi_0(-ik_1(0), \mu)\, d\mu = 1$). Note that the representation of Eq. (7.70) in some other form has been also used by Ivanov and Volkov (1978); Rogovtsov and Samson (1985a). Now use the representation of Eq. (7.70) to reduce the IE of Eq. (7.66) to the following form:

$$\begin{aligned} & R^0(|\mu|, \mu_1, \omega_0) \\ &= \frac{1}{2}\tilde{G}_{\infty;0}(+0, -|\mu|; 0, \mu_1; \omega_0) - c_1(\omega_0; 0)\, \beta(\mu_1; \omega_0)\, \Phi_0(-ik_1(0), -|\mu|) \\ &\quad - \int_0^1 \mu' \tilde{G}^*_{\infty;0}(+0, -|\mu|; 0, -\mu'; \omega_0)\, R^0(\mu', \mu_1, \omega_0)\, d\mu', \end{aligned} \tag{7.71}$$

where $\beta(\mu_1; \omega_0) = \int_0^1 \mu' \Phi_0(-ik_1(0), -\mu')\, R^0(\mu', \mu_1, \omega_0)\, d\mu'$. Equation (7.71) can also be regarded as the integral equation with respect to function $R^0(|\mu|, \mu_1, \omega_0)$. As first shown by Rogovtsov and Samson (1976), such an equation can be most effectively solved, if one first treats the value $\beta(\mu_1; \omega_0)$ as known and then finds it

just after the formal solution to these integral equations. Using such a technique of the solution, the following linear integral operator plays an important role:

$$\hat{K}\left(f\right)\left(\mu\right)=-\int_{0}^{1}K\left(\mu,\mu'\right)f\left(\mu'\right)\,d\mu',\ \ \mu\in\left[0,1\right], \tag{7.72}$$

where $K\left(\mu,\mu'\right)$ stands for function $Q_1\left(\left|\mu\right|,\left|\mu'\right|,\omega_0\right)=\left(\left|\mu'\right|\tilde{G}^{*}_{\infty;0}(+0,-\left|\mu\right|;0,-\left|\mu'\right|;\omega_0)\right)$. Let $\hat{K}^l$ denote the lth power of operator $\hat{K}$. This power is defined by the relation of $\hat{K}^l\left(f\right)\left(\mu\right)=\hat{K}\left(\hat{K}^{l-1}\left(f\right)\right)\left(\mu\right)$, where $l\in\{1,2,3,\dots\}$ and $\mu\in[0,1]$. Symbol $\hat{K}^0$ is understood as the identity operator (i.e. $\hat{K}^0\left(f\right)\left(\mu\right)=f\left(\mu\right)$). Let the integral operator $\hat{K}$ be a contraction operator (Riesz and Sz.-Nagy, 1972; Richtmyer, 1978; Yosida, 1980) in some normalized space X of functions defined on segment [0, 1]. This means that, for all functions $f\left(\mu\right)\in X$, an inequality $||\hat{K}\left(f\right)||_X\leq q||\left(f\right)||_X$ holds, where q is some positive constant less than unity (by symbol $||\dots||_X$ is meant the norm in space X). Taking into account the definition of integral operator $\hat{K}$ and of quantity $\beta\left(\mu_1;\omega_0\right)$, and the principle of contracting mappings (see, for example, Edwards (1965)), the formal solution to the IE of Eq. (7.71) in analytical form can be written down as follows:

$$\begin{aligned}R^0\left(\left|\mu\right|,\mu_1,\omega_0\right)\ &=\ 2^{-1}\Upsilon^{-}_{1;tot}\left(\left|\mu\right|,\mu_1,\omega_0\right)-2^{-1}c_1\left(\omega_0;0\right)\Upsilon^{-}_{1}\left(\left|\mu\right|,\omega_0\right)\\&\times\left(\int_0^1\mu'\Phi_0\left(-ik_1\left(0\right),-\mu'\right)\Upsilon^{-}_{1;tot}\left(\mu',\mu_1,\omega_0\right)d\mu'\right)\\&\times\left(1+c_1\left(\omega_0;0\right)\int_0^1\mu'\Phi_0\left(-ik_1\left(0\right),-\mu'\right)\Upsilon^{-}_{1}\left(\mu',\omega_0\right)d\mu'\right)^{-1}.\end{aligned} \tag{7.73}$$

Functions $\Upsilon^{-}_{1;tot}\left(\left|\mu\right|,\mu_1,\omega_0\right)$ and $\Upsilon^{-}_{1}\left(\left|\mu\right|,\omega_0\right)$ in Eq. (7.73) can be calculated by the iteration method to be formally defined by expressions $\Upsilon^{-}_{1;tot}\left(\left|\mu\right|,\mu_1,\omega_0\right)=\sum_{r=0}^{+\infty}\hat{K}^r(\tilde{G}_{\infty;0})\left(\left|\mu\right|\right)$ and $\Upsilon^{-}_{1}\left(\left|\mu\right|,\omega_0\right)=\sum_{r=0}^{+\infty}\hat{K}^r(\Phi_{0;1})\left(\left|\mu\right|\right)$. Here, symbols $\tilde{G}_{\infty;0}$ and $\Phi_{0;1}$ should be understood as functions $\tilde{G}_{\infty;0}\left(+0,-\left|\mu\right|;0,\mu_1;\omega_0\right)$ and $\Phi_0\left(-ik_1\left(0\right),-\left|\mu\right|\right)$, respectively.

The presentation of the solution to the IE of Eq. (7.71) in the form of Eq. (7.73) is possible, if there holds the inequality of

$$\sup_{\mu_1\in(0,1]}\left\|Q_1\right\|_{L_1(0,1)}=\sup_{\mu_1\in(0,1]}\left[\int_0^1\left|Q_1\left(\mu,\mu_1,\omega_0\right)\right|\,d\mu\right]\leq q<1$$

(or $\left\|Q_1\right\|_{L_2(0,1)}=\left[\int_0^1 d\mu\int_0^1\left(Q_1\left(\mu,\mu_1,\omega_0\right)\right)^2d\mu_1\right]^{\frac{1}{2}}\leq q_1<1$). The term $\sup_{\mu_1\in(0,1]}\left[f\left(\mu_1\right)\right]$ is for the least upper bound of the set of values, which are accepted by function $f\left(\mu_1\right)$ on half-interval $(0,1]$). When the above inequality is valid, there will exist a unique solution to the IE of Eq. (7.71) within the class of absolutely integrable functions on [0, 1] (i.e. in the class of $L_1\left(0,1\right)$ or the class of square summable functions on [0, 1] (i.e. in the class of $L_2\left(0,1\right)$). Also, this solution can be represented in the form of Eq. (7.73). The fulfillment of these inequalities

should be verified for each phase function selected. Note that the general physical considerations and specific calculations indicate that $\sup_{\mu_1\in(0,1]}\|Q_1\|_{L_1(0,1)}$ (or $\|Q_1\|_{L_2(0,1)}$) assumes values, which are about an order of magnitude less (in decimal notation) than unity for the cases of isotropic and forward-extended geophysical phase functions. This fact has also been noted in Rogovtsov and Samson (1976) for the case of isotropic scattering. Besides, there has been shown by Rogovtsov and Borovik (2009) that, for the case of the water cloud C.1 model ($\lambda = 0.6457\mu$m; *lambda* is the wavelength), value $\sup_{\mu_1\in(0,1]}\|Q_1\|_{L_1(0,1)}$ is approximately 0.1055 (for this model the single scattering albedo ω_0 is equal to 1). If ω_0 is less than unity, then $\sup_{\mu_1\in(0,1]}\|Q_1\|_{L_1(0,1)}$ will be further less. Table 7.1 gives, for illustration, the estimations of $\sup_{\mu_1\in(0,1]}\|Q_1\|_{L_1(0,1)}$ for the case of the Henyey–Greenstein phase function with $g = f_1 = 0.9$.

Table 7.1. Estimations of value $\sup_{\mu_1\in(0,1]}\|Q_1\|_{L_1(0,1)}$ for various values ω_0 for the case of the Henyey–Greenstein phase function with parameter $g = 0.9$

ω_0	0.1	0.2	0.3	0.4	0.5
$\|Q_1\|_{L_1(0,1)}$	$0.1466\cdot10^{-3}$	$0.6461\cdot10^{-3}$	$0.1619\cdot10^{-2}$	$0.3245\cdot10^{-2}$	$0.5813\cdot10^{-2}$
ω_0	0.6	0.7	0.8	0.9	0.99
$\|Q_1\|_{L_1(0,1)}$	$0.9899\cdot10^{-2}$	$0.1619\cdot10^{-1}$	$0.2696\cdot10^{-1}$	$0.4778\cdot10^{-1}$	$0.9410\cdot10^{-1}$
ω_0	0.999	0.9999	0.99999	0.999999	
$\|Q_1\|_{L_1(0,1)}$	0.1026	0.10353	0.10363	0.10364	

Table 7.2 gives the nonnegative smallest element $k_1(0)$ of the discrete spectrum of the characteristic equation of the RTT and the normalization constants for various values of single scattering albedo ω_0 for the case of the Henyey–Greenstein phase function with parameter $g = 0.9$.

Table 7.2. Values of $k_1(0)$ and normalization constants $c_1(\varpi_0;0)$

ω_0	0.1	0.2	0.3	0.4	0.5
$k_1\,(0)$	0.94971	0.87827	0.79928	0.71489	0.62576
$c_1\,(\omega_0;0)$	$0.75198\cdot10^{-1}$	0.13989	0.20717	0.28291	0.37362
ω_0	0.6	0.7	0.8	0.9	0.99
$k_1\,(0)$	0.53185	0.43247	0.32580	0.20636	$0.56047\cdot10^{-1}$
$c_1\,(\omega_0;0)$	0.48981	0.65240	0.91433	$0.14815\cdot10^{1}$	$0.53808\cdot10^{1}$
ω_0	0.999	0.9999	0.99999	0.999999	0.9999999
$k_1\,(0)$	$0.17362\cdot10^{-1}$	$0.54785\cdot10^{-2}$	$0.17321\cdot10^{-2}$	$0.54772\cdot10^{-3}$	$0.17321\cdot10^{-3}$
$c_1\,(\omega_0;0)$	$0.17289\cdot10^{2}$	$0.54762\cdot10^{2}$	$0.17320\cdot10^{3}$	$0.54772\cdot10^{3}$	$0.17321\cdot10^{4}$

It follows from what has been said above and from Table 7.1 that the iteration procedures, which should be utilized for deriving functions $\Upsilon^{-}_{1;tot}\,(|\mu|,\mu_1,\omega_0)$ and $\Upsilon^{-}_{1}\,(|\mu|,\omega_0)$, will be very effective. Their efficiency was demonstrated, in particular, by Rogovtsov and Borovik (2009) for the case of conservative scattering.

While deriving functions $\Upsilon^{-}_{1;tot}(|\mu|,\mu_1,\omega_0)$ and $\Upsilon^{-}_{1}(|\mu|,\omega_0)$ by means of the iteration procedures, functions $\tilde{G}_{\infty;0}(+0,-|\mu|;0,\mu_1;\omega_0)$ and $\Phi_0(-ik_1(0),-|\mu|)$ should be taken as initial ones. The analytical expressions and the algorithms for finding these functions have been obtained and described in Rogovtsov (1996) and Rogovtsov and Borovik (2009). So, we will succeed in the effective reduction of the problem on deriving the function $R^0(|\mu|,\mu_1,\omega_0)$ to the search of particular values of the Green function for the RTE for a plane-parallel medium infinite turbid medium V_∞ by using the reduction procedure of one of the general invariance relations. In its turn, the analytical representations of this Green function have been obtained by Rogovtsov and Borovik (1993, 2009).

It follows from the above and from Table 7.1 that, for obtaining rather exact approximate analytical solutions for function $R^0(|\mu|,\mu_1,\omega_0)$, it is enough to replace functions $\Upsilon^{-}_{1;tot}(|\mu|,\mu_1,\omega_0)$ and $\Upsilon^{-}_{1}(|\mu|,\omega_0)$ by functions $\tilde{G}_{\infty;0}(+0,-|\mu|;0,\mu_1;\omega_0)$ and $\Phi_0(-ik_1(0),-|\mu|)$, respectively (here, however, inequality $1-\omega_0 \ll 1$ should not be valid).

Table 7.3 gives spherical albedo $A_s(\omega_0)$ values for various ω_0 for the case of Henyey–Greenstein phase function with parameter $g = 0.9$. These values were obtained with using Eq. (7.73), the definition of $A_s(\omega_0)$ (see, for example, Sobolev, 1975), and algorithms for calculating functions $Q_1(|\mu|,\mu_1,\omega_0)$and $\tilde{G}_{\infty;0}(+0,-|\mu|;0,\mu_1;\omega_0)$ described by Rogovtsov (1996) and Rogovtsov and Borovik (1993, 2009).

Table 7.3. Values of spherical albedo $A_s(\omega_0)$

ω_0	0.1	0.2	0.3	0.4	0.5
A_s	$0.20561\cdot10^{-2}$	$0.46210\cdot10^{-2}$	$0.79046\cdot10^{-2}$	$0.12249\cdot10^{-1}$	$0.18254\cdot10^{-1}$
ω_0	0.6	0.7	0.8	0.9	0.99
A_s	$0.35520\cdot10^{-1}$	$0.41226\cdot10^{-1}$	$0.67430\cdot10^{-1}$	0.13259	0.49342
ω_0	0.999	0.9999	0.99999	0.999999	
A_s	0.79543	0.92975	0.97719	0.99273	

The algorithm presented in this subsection can be used even for the case of phase functions highly extended in the forward direction (to approximate such a phase function it is necessary to use several thousands of addends in its expansion in Legendre polynomials). It should be noted that the algorithms presented in this subsection and in the literatur (Loyalka, 1974; Rogovtsov and Samson, 1976; Domke, 1976; Ganapol and Kornreich, 2005; Rogovtsov and Borovik, 2009) have a number of general features. All algorithms presented there use Fredholm equations that have been obtained in essence with the help of related procedures and can be solved with the help of effective iterative methods. In addition, these algorithms directly use (in explicit or implicit form) relations connecting Green's functions (or functions related to Green's functions) of the radiative (neutron) transfer equation for the case of infinite plane-parallel medium with solutions of initial boundary-value problems for the cases of semi-infinite or finite plane-parallel layers.

7.4.2 Asymptotic and approximate analytical expressions for monochromatic radiation fluxes exiting macroscopically homogeneous non-concave turbid bodies

Assume that all the conditions used for deriving the two-sided inequalities of Eq. (7.48) are fulfilled. Then monochromatic radiance $F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$ of body V_β can be approximately calculated by the following formula (Rogovtsov, 1983, 1994):

$$F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \approx \left[1-2^{-1}\left(b\left(\mathrm{S}_\beta\right)+a\left(\mathrm{S}_\beta\right)\right)\right] \times \left[\left(1-b\left(\mathrm{S}_\beta\right)\right)\left(1-a\left(\mathrm{S}_\beta\right)\right)\right]^{-1} F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_{\beta_4}\right). \quad (7.74)$$

At this, the relative error of this relation will not exceed the quantity $e_0 = \left(2\left(1-a\left(\mathrm{S}_\beta\right)\right)\right)^{-1}\left(a\left(\mathrm{S}_\beta\right)-b\left(\mathrm{S}_\beta\right)\right)$. Equation (7.74) is most interesting for the case when body V_{β_4} coincides with macroscopically homogeneous infinite turbid medium V_∞. This is due to the fact that quantities $F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_{\beta_4}\right)$, $b\left(\mathrm{S}_\beta\right)$, and $a\left(\mathrm{S}_\beta\right)$ in this case can be found in an explicit form or, at least, be estimated.

Usage of the two-sided inequalities of Eq. (7.49) enables quantity $F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$ to be estimated with the following approximate simple formula (Rogovtsov, 1994):

$$F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \approx \left(4-\omega_0\right)\left(2\left(2-\omega_0\right)\right)^{-1} F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_{\beta_4}\right). \quad (7.75)$$

The relative error of Eq. (7.75) does not exceed $e_0 = \left(2\left(2-\omega_0\right)\right)^{-1}\omega_0$ ($e_0 \le 2^{-1}$ for all $\omega_0 \in (0,1)$). Equation (7.75) can only be used for the case of macroscopically homogeneous turbid bodies and phase functions that are constant or forward-extended.

Equation (7.74) can be effectively utilized to find quantity $F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$ in situations, at least, when $\mathrm{V}_{\beta_4} = \mathrm{V}_\infty$ and a macroscopically homogeneous non-concave turbid body V_β is bounded by a smooth surface S_β. Moreover, Eq. (7.34) is an asymptotic formula for monochromatic radiance $F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$ for the case of nearly conservatively-scattering optically-thick medium. Such a turbid medium stands for body V_β, for which two conditions are fulfilled. First, the minimal positive element $k_1(0)$ of the discrete spectrum of the characteristic equation of the RTT (see, for example, Rogovtsov (1996) and Rogovtsov and Borovik (2009) corresponding to the zeroth azimuthal harmonics should satisfy the inequality of $\upsilon = \left(k_1(0)\right)\left(3\left(1-f_1\right)\right)^{-1} \ll 1$ (i.e. $\upsilon_0 = \sqrt{\left(1-\omega_0\right)\left(3\left(1-f_1\right)\right)^{-1}} \ll 1$). Second, the inequality of $k_1(0)\,\kappa_{ext} l\left(\mathrm{S}_\beta\right) \gg 1$ should hold, where $l\left(\mathrm{S}_\beta\right)$ is the greatest lower bound of the set of all values of the curvature radii corresponding to the points belonging to boundary S_β of body V_β. Write down the asymptotic formula for quantity $F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$ for the case of the turbid medium with the above-said properties. To do so, introduce first quantity $\tau_0\left(\vec{r}_{\mathrm{S}_\beta};\mathrm{S}_\beta\right)$, which has the sense of the least upper bound of the set of optical radius values of spheres that are tangent to surface S_β at a point specified by radius-vector $\vec{r}_{\mathrm{S}_\beta}$ and are fully contained in body V_β (the optical lengths in macroscopically homogeneous turbid media, which differ from a slab in shape, everywhere below stand for the products of κ_{ext} by the geometrical lengths). In its turn, denote the greatest lower bound of the set of values $\tau_0\left(\vec{r}_{\mathrm{S}_\beta};\mathrm{S}_\beta\right)$ that it assumes, when the end of radius-vector $\vec{r}_{\mathrm{S}_\beta}$ goes through all the points of surface S_β, via $\tau_0\left(\mathrm{S}_\beta\right)$. There has been shown by Rogovtsov (1988a,

1991c, 1994) that, for the case of nearly conservatively-scattering, optically-thick, macroscopically homogeneous turbid medium V_β and the limited phase function, the following asymptotic formulae are valid:

$$\begin{aligned} a\left(\mathrm{S}_\beta\right) &= 2^{-1} + \mathrm{O}\left(k_1\left(0\right) + \left(k_1\left(0\right)\tau_0\left(\mathrm{S}_\beta\right)\right)^{-1}\right), \\ b\left(\mathrm{S}_\beta\right) &= 2^{-1} + \mathrm{O}\left(k_1\left(0\right) + \left(k_1\left(0\right)\tau_0\left(\mathrm{S}_\beta\right)\right)^{-1}\right), \end{aligned} \tag{7.76}$$

where $\upsilon \to 0$, $k_1\left(0\right)\tau_0\left(\mathrm{S}_\beta\right) \to +\infty$.

Within the frame of the accepted assumptions, one obtains the asymptotic formula below from Eqs (7.34) and (7.76) (Rogovtsov, 1988a, 1991c, 1994):

$$F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \sim 2F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right),\ \upsilon \to 0,\ k_1\left(0\right)\tau_0\left(\mathrm{S}_\beta\right) \to +\infty. \tag{7.77}$$

It follows from the above relation for the e_0 value, which gives the upper estimation of the *a priori* relative error of Eq. (7.74), that the relative error of Eq. (7.77) has the order of the residual terms of Eq. (7.76). It should be especially noted that the asymptotic formula of Eq. (7.77) is one of the few analytical solutions to boundary-value problems for the RTT for the case of non-concave macroscopically homogeneous turbid bodies, the shape of which can have no symmetry at all. This conclusion becomes more convincing, if one takes into account that monochromatic flux $F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)$ through surface S_β of body V_β comprising arbitrary internal radiation sources and being imbedded together with the sources in infinite medium V_∞ can be analytically found. The corresponding formula for $F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$ will be given below.

If turbid medium V_β is a plane-parallel layer of optical thickness τ_0, then more substantial asymptotics formulae, as compared with Eq. (7.76), can be derived for quantities $b\left(\mathrm{S}_\beta\right)$ and $a\left(\mathrm{S}_\beta\right)$ for the case of nearly conservative scattering. These asymptotic formulae have been obtained by using the representation of Eq. (7.70) and asymptotic expressions for quantities $c_1\left(\omega_0, 0\right)$ and $\Phi_0\left(-ik_1\left(0\right), \mu\right)$ (see, for example, the publications (Sobolev, 1975; Yanovitskij, 1997; Rogovtsov and Samson, 1985a; Rogovtsov, 1991b)). The said asymptotic formulae have the following form (Rogovtsov, 1986b, 1994):

$$\begin{aligned} b\left(\mathrm{S}_\beta\right) &= 2^{-1}\left(1 - \exp\left(-k_1\left(0\right)\tau_0\right)\right)\left(1 - 3\upsilon_0\right) + O(\upsilon_0^2), \\ a\left(\mathrm{S}_\beta\right) &= 2^{-1}\left(1 - \exp\left(-k_1\left(0\right)\tau_0\right)\right) + O(\upsilon_0^2),\ \upsilon_0 \to 0. \end{aligned} \tag{7.78}$$

The formulae of Eqs (7.78) are valid for any values $\tau_0 > 0$ and any phase functions. Although value $k_1\left(0\right)$, for rather small values of $\left(1 - \omega_0\right)$, will be too small (for $k_1\left(0\right)$, there holds the asymptotic formula of $k_1\left(0\right) \sim \sqrt{3\left(1 - f_1\right)\left(1 - \omega_0\right)}$ at $\upsilon_0 \to 0$; see, for example, Sobolev (1975)), product $k_1\left(0\right)\tau_0$ can assume, in principle, any positive values at rather large values of optical thickness τ_0 of layer V_β. Due to the above-said, quantities $b\left(\mathrm{S}_\beta\right)$ and $a\left(\mathrm{S}_\beta\right)$ will approach $2^{-1}\left(1 - \exp\left(-k_1\left(0\right)\tau_0\right)\right)$ at $\upsilon_0 \to 0$. Using Eqs (7.74) and (7.78), it can be shown that the following asymptotic formula is valid for any $\tau_0 > 0$:

$$\begin{aligned} F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) &\sim 2\left(1 + \exp\left(-k_1\left(0\right)\tau_0\right)\right)^{-1} \\ &\quad \times \left[1 - 2^{-1}3\upsilon_0 \tanh\left(2^{-1}k_1\left(0\right)\tau_0\right)\right] F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right), \upsilon_0 \to 0. \end{aligned} \tag{7.79}$$

Moreover, the following asymptotic formula holds for upper estimation e_0^* of the relative error of Eq. (7.79):

$$e_0^* = 3\upsilon_0 2^{-1}\tanh\left(2^{-1}k_1(0)\,\tau_0\right) + \mathrm{O}(\upsilon_0^2),\ \upsilon_0 \to 0. \tag{7.80}$$

While deriving Eqs (7.74), (7.75), (7.79), and (7.80), it was assumed that the monochromatic power of all internal stationary radiation sources in turbid medium V_β is a finite quantity (i.e. the inequality of $\mathrm{P}\left(\mathrm{V}_\beta^0\right) = \iiint\limits_{\mathrm{V}_\beta^0} dV \int_\Omega g_\beta\left(\vec{r},\vec{\Omega};\mathrm{V}_\beta^0\right) d\Omega < +\infty$ holds). If body V_β has the shape of a plane slab comprising internal sources with the plane-parallel symmetry yet, then one can refuse this limitation. Let $g_\beta\left(z,\vec{\Omega};\mathrm{V}_\beta^0\right)$ be a density of internal sources in such a slab V_β (z is the shortest distance from an 'observation' point to its first boundary; assume, for the unambiguity, that the Z-axis is directed from its first boundary to the second one and is orthogonal to them). Then all Eqs (7.74), (7.75), (7.79), and (7.80) are valid. However, quantity $F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$ of Eqs (7.74), (7.75), and (7.79) should be understood as the monochromatic radiation power exiting the slab through the end faces of cylinder V_{β^*} with the unit section (the said end faces should lie on the corresponding boundaries of the plane-parallel slab). Respectively, quantity $F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)$ of the said formulae should be understood as the sum of the radiation fluxes through the end faces of the above cylinder, which lies inside slab V_β and, together with it, is embedded into macroscopically homogeneous infinite medium V_∞ (while calculating the radiation fluxes through each end face, one should consider the external unit normals to them).

Assume that quantity $\mathrm{P}\left(\mathrm{V}_\beta^0\right)$ is a finite one ($\mathrm{P}\left(\mathrm{V}_\beta^0\right)$ is the monochromatic power of all radiation sources in V_β^0). Then the difference of $\mathrm{E}_a\left(\mathrm{V}_\beta^0\right) = \left(\mathrm{P}\left(\mathrm{V}_\beta^0\right) - F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)\right)$ has the sense of the radiation energy that is absorbed per unit time in unit frequency range by the whole turbid body V_β. If body V_β is bounded by a smooth surface and is a nearly conservatively scattering and optically thick medium, then one gets the following asymptotic formula taking account of Eq. (7.77):

$$\mathrm{E}_a\left(\mathrm{V}_\beta^0\right) \sim \iiint\limits_{\mathrm{V}_\beta^0} dV \int\limits_\Omega g_\beta\left(\vec{r},\vec{\Omega};\mathrm{V}_\beta^0\right) d\Omega - 2F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right),$$

$$k_1(0)\left(3\left(1-f_1\right)\right)^{-1} \to 0,\ k_1(0)\ \tau_0\left(\mathrm{S}_\beta\right) \to +\infty. \tag{7.81}$$

If V_β is a plane slab with optical thickness τ_0 and $\mathrm{P}\left(\mathrm{V}_\beta^0\right) < +\infty$, then one finds the following asymptotic formula taking account of Eq. (7.79):

$$\mathrm{E}_a\left(\mathrm{V}_\beta^0\right) \sim \iiint\limits_{\mathrm{V}_\beta^0} dV \int_\Omega g_\beta\left(\vec{r},\vec{\Omega};\mathrm{V}_\beta^0\right) d\Omega - \frac{2F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)}{1+\exp\left(-k_1(0)\,\tau_0\right)}$$
$$\times\left[1-\frac{3}{2}\upsilon_0\tanh\left(\frac{k_1(0)\,\tau_0}{2}\right)\right],\ \upsilon_0 \to 0. \tag{7.82}$$

Now let V_β be a slab comprising internal sources described by density $g_\beta\left(z,\vec{\Omega};\mathrm{V}_\beta^0\right)$ (i.e. the internal sources possess the plane-parallel symmetry). Then $\mathrm{P}\left(\mathrm{V}_\beta^0\right) = +\infty$. Therefore, Eq. (7.84) should be replaced by the following asymptotic formula:

$$\mathrm{E}_a\left(\mathrm{V}^0_{\beta^*}\right) \sim \int_0^{z_0} dz \int_\Omega g_\beta\left(z,\vec{\Omega};\mathrm{V}^0_\beta\right) d\Omega - \frac{2F_2\left(S_{\beta^*}\left(\mathrm{V}^0_{\beta^*}\right);\mathrm{V}_\infty\right)}{1+\exp\left(-k_1\left(0\right)\tau_0\right)} \times \left[1-\frac{3}{2}\upsilon_0 \tanh\left(\frac{k_1\left(0\right)\tau_0}{2}\right)\right], \upsilon_0 \to 0. \tag{7.83}$$

Here $F_2\left(S_{\beta^*}\left(\mathrm{V}^0_{\beta^*}\right);\mathrm{V}_\infty\right)$ is the sum of the radiation fluxes through the end faces of cylinder V_{β^*} (at this, the end faces lie on the respective boundaries of slab V_β that is embedded into medium V_∞); z_0 is the geometrical thickness of slab V_β. Some particular consequences (they refer to the case of the illumination of the slab V_β by external sources) have been derived earlier by Rogovtsov (1991b, 1994); Yanovitskij (1991).

Equations (7.74), (7.75), (7.77), (7.79), and (7.81)–(7.83) can be constructively utilized, if one knows an effective method to find both the monochromatic flux $F_2\left(\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right);\mathrm{V}_{\beta_4}\right)$ through the whole boundary S_β of body V_β and the sum of the radiation fluxes through the end faces of cylinder V_{β^*}, i.e. quantity $F_2\left(S_{\beta^*}\left(\mathrm{V}^0_{\beta^*}\right);\mathrm{V}_\infty\right)$. Such a method has been proposed by Rogovtsov (1988a, 1991b, 1991c, 1994, 1997) for the case of macroscopically homogeneous body V_β. This method is based on using the following expression:

$$F_2\left(\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right);\mathrm{V}_\infty\right) = \iiint\limits_{\mathrm{V}^0_\beta} dV \int_\Omega \left(1-\mathrm{I}^{\times\times}\left(\vec{r},-\vec{\Omega};\mathrm{S}_\beta;\mathrm{V}_\infty\right)\right) g_\beta\left(\vec{r},\vec{\Omega};\mathrm{V}^0_\beta\right) d\Omega, \tag{7.84}$$

where $\mathrm{I}^{\times\times}\left(\vec{r},-\vec{\Omega};\mathrm{S}_\beta;\mathrm{V}_\infty\right)$ is defined by Eq. (7.47), where one should assume that $m\left(\vec{r}\right) \equiv m = \mathrm{const}$, $\kappa_{ext}\left(\vec{r}\right) \equiv \kappa_{ext} = \mathrm{const}$, $\kappa_{sca}\left(\vec{r}\right) \equiv \kappa_{sca} = \mathrm{const}$, and $S_{\beta_2} = \mathrm{S}_\beta$. Equation (7.84) has been derived on the basis of using the definition for quantity $F_2\left(\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right);\mathrm{V}_\infty\right)$ (see Eq. (7.46)) and of the stationary analog of Eq. (7.74). Function $\mathrm{I}^{\times\times}\left(\vec{r},\vec{\Omega};\mathrm{S}_\beta;\mathrm{V}_\infty\right)$ can be represented as follows for the case of a macroscopically homogeneous medium:

$$\mathrm{I}^{\times\times}\left(\vec{r},\vec{\Omega};\mathrm{S}_\beta;\mathrm{V}_\infty\right) = \left(\kappa_{ext}-\kappa_{sca}\right) \iiint\limits_{\mathrm{V}^0_\beta} G_*\left(\vec{r},\vec{\Omega};\vec{r}\,';\mathrm{V}_\infty\right) dV', \tag{7.85}$$

where $G_*\left(\vec{r},\vec{\Omega};\vec{r}\,';\mathrm{V}_\infty\right)$ is the Green function of Eq. (7.10) for the case of infinite macroscopically homogeneous medium V_∞, which comprises a point isotropic source with density $\delta\left(\vec{r}-\vec{r}\,'\right)$. Due to the spherical symmetry inherent in the radiation field produced by such a 'source' in medium V_∞, this Green function can be written as $G_*\left(\vec{r},\vec{\Omega};\vec{r}\,';\mathrm{V}_\infty\right) = G_*\left(\left|\vec{r}-\vec{r}\,'\right|,\left(\vec{\rho}\cdot\vec{\Omega}\right);\mathrm{V}_\infty\right)$, where $\vec{\rho} = \left(\left|\vec{r}-\vec{r}\,'\right|\right)^{-1}\left(\vec{r}-\vec{r}\,'\right)$. Various analytical representations of function $\tilde{G}_*\left(\left|\vec{\tau}-\vec{\tau}\,'\right|,\mu;\tilde{\mathrm{V}}_\infty\right) = \left(\kappa_{ext}\right)^{-2} G_*\left(\left|\vec{r}-\vec{r}\,'\right|,\left(\vec{\rho}\cdot\vec{\Omega}\right);\mathrm{V}_\infty\right)$ have been obtained by Kolesov (1983), Rogovtsov (1986a, 1988b) and Freimanis (2005). Quantities $\vec{\tau}$, $\vec{\tau}\,'$ and μ are stood for $\vec{\tau} = \kappa_{ext}\vec{r}$, $\vec{\tau}\,' = \kappa_{ext}\vec{r}\,'$, and $\mu = \left(\left|\vec{\tau}-\vec{\tau}\,'\right|\right)^{-1}\left(\left|\vec{\tau}-\vec{\tau}\,'\right|\cdot\vec{\Omega}\right)$, respectively, and $\tilde{\mathrm{V}}_\infty$ is stood for the image of body V_∞ under the use of the transformation of $\vec{r} \to \kappa_{ext}\vec{r}$. Using the analytical representations of the above-said works and Eqs (7.84) and (7.85), one can get the analytical relations for quantity $F_2\left(\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right);\mathrm{V}_\infty\right)$. The simplest analytical expressions for $F_2\left(\mathrm{S}_\beta\left(\mathrm{V}^0_\beta\right);\mathrm{V}_\infty\right)$

have been obtained in the papers (Rogovtsov, 1988a, 1994, 1997) within the scope of the assumptions on the isotropy of internal radiation sources in V_β^0 (i.e. $g_\beta\big(\vec{r},\vec{\Omega};\mathrm{V}_\beta\big) \equiv g_\beta\big(\vec{r};\mathrm{V}_\beta^0\big)$). In particular, quantity $F_2\big(\mathrm{S}_\beta\big(\mathrm{V}_\beta^0\big);\mathrm{V}_\infty\big)$ can be represented as follows (Rogovtsov, 1994):

$$F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right) = \frac{1-\varpi_0}{(\kappa_{ext})^3} \iiint\limits_{\tilde{\mathrm{V}}_\beta^0} g_\beta\big(\vec{\tau}';\tilde{\mathrm{V}}_\beta^0\big)d\vec{V} \iiint\limits_{\tilde{\mathrm{V}}_\infty\setminus\tilde{\mathrm{V}}_\beta} \tilde{G}_*\big(\vec{\tau}',\vec{\tau};\vec{\mathrm{V}}_\infty\big)\,d\tilde{V}. \quad (7.86)$$

Here $\tilde{g}_\beta\big(\vec{\tau}';\tilde{\mathrm{V}}_\beta^0\big) = g_\beta\big((\kappa_{ext})^{-1}\vec{\tau}';\mathrm{V}_\beta^0\big)$, $\tilde{\mathrm{V}}_\beta^0$ is an image of body V_β^0 under the transformation of $\vec{r} \to \kappa_{ext}\vec{r}$, and function $\tilde{G}_*\big(\vec{\tau}',\vec{\tau};\tilde{\mathrm{V}}_\infty\big)$ is defined by the following relation:

$$\begin{aligned}\tilde{G}_*\big(\vec{\tau}',\vec{\tau};\tilde{\mathrm{V}}_\infty\big) &= -\left(|\vec{\tau}'-\vec{\tau}|\right)^{-1}\frac{\partial}{\partial\tau^*}\left[\int_{-1}^{1}d\mu\int_{-1}^{1}\tilde{G}_{\infty;0}\left(\tau^*,\mu;0,\mu';\varpi_0\right)d\mu'\right]\Bigg|_{\tau^*=|\vec{\tau}'-\vec{\tau}|}\\ &= \tilde{G}_*\big(\,|\vec{\tau}'-\vec{\tau}|\,;\tilde{\mathrm{V}}_\infty\big). \end{aligned} \quad (7.87)$$

Equation (7.87) comprises the Green function $\tilde{G}_{\infty;0}\left(\ldots\right)$ for the dimensionless RTE (7.67) for the case of an infinite homogeneous plane-parallel medium $\tilde{\mathrm{V}}_\infty$ containing a 'source' with density $\delta\left(\tau\right)\delta\left(\mu-\mu'\right)$. A number of analytical representations of this function have been obtained by Rogovtsov and Borovik (1993, 2009) for the case of an arbitrary phase function (in particular, for forward extended phase functions). The simplest expression for the function in the square brackets of Eq. (7.87) has been derived in the papers (Rogovtsov, 1994, 1997). It has the following form:

$$\begin{aligned}&\int_{-1}^{1}d\mu\int_{-1}^{1}\tilde{G}_{\infty;0}\left(\tau,\mu;0,\mu';\omega_0\right)\,d\mu'\\ &= \int_{1}^{+\infty}s^{-1}\exp\left(-s\tau\right)\,ds\\ &\quad+\frac{1}{2\pi}\int_{-\infty}^{+\infty}\left(2\mathfrak{A}_0\left(\alpha^2\right)-\mathfrak{M}\left(\alpha\right)\right)\exp\left(-i\alpha\tau\right)\,d\alpha = \tilde{G}_{\infty;0}\left(\tau,\omega_0\right). \end{aligned} \quad (7.88)$$

Here $\tau>0$, $\mathfrak{A}_0\left(\alpha^2\right) = \left(\left(1-\omega_0\right)\wp\left(\alpha^2\right)\right)^{-1}$; $\wp\left(\alpha^2\right)$ is the infinite continued fraction that is defined by the following relation:

$$\wp\left(\alpha^2\right) = 1+\cfrac{q_0\left(0\right)\alpha^2}{1+\cfrac{q_1\left(0\right)\alpha^2}{1+\ldots}}; \quad (7.89)$$

and function $\mathfrak{M}\left(\alpha\right)$ is equal to $\int_{-1}^{1}\left(1-i\alpha\mu\right)^{-1}d\mu$. The definitions of quantities $q_l\left(0\right)$ for any $l=0,\ 1,\ 2,\ \ldots$ are given in the explanations to Eq. (7.70). It is important to emphasize that the integrand function of the second integral of the right-hand side of Eq. (7.88) is a limited function for any $\omega_0\in(0,1)$, any $\alpha\in(-\infty,+\infty)$, and any $\tau\in(-\infty,+\infty)$. Besides, there exists the asymptotic formula of $2\mathfrak{A}_0\left(\alpha^2\right) - \mathfrak{M}\left(\alpha\right) = O\left(\alpha^{-2}\right)$ for $|\alpha|\to+\infty$. This fact is useful to be accounted for, while calculating the second integral of the right hand of Eq. (7.88).

The contour integration method and the information on the analytical properties of infinite continued fraction $\wp_0\left(\alpha^2\right)$ (see, for example, Rogovtsov (1994, 1996) and Rogovtsov and Borovik (2009) can be also used to find this integral for the case of an arbitrary phase function. If V_β is a plane-parallel turbid medium, then quantity $F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)$ can be calculated for any function $g_\beta\left(z,\vec{\Omega};\mathrm{V}_\beta^0\right)$ and for any phase function by means of the general analytical expressions derived in the papers (Rogovtsov, 1991b, 1994) and of the analytical representations of the Green function $\tilde{G}_{\infty;0}\left(\tau,\mu;0,\mu';\omega_0\right)$. Various analytical formulae for this Green function were derived by Rogovtsov and Borovik (1993, 2009).

To illustrate the opportunities of the GIRRM in obtaining rather simple analytical solutions to boundary-value problems of the RTT, let us give a number of asymptotic formulae for monochromatic luminosity $F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$ and for quantity $F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)$ that have been deduced in the papers (Rogovtsov, 1988a, 1994, 1997). Let macroscopically homogeneous turbid medium V_β (body V_β can generally have a concave shape) contain an internal point isotropic source with power $P\left(\mathrm{V}_\beta^0\right)$. Hence, the radiation density corresponding to this source is $g_\beta\left(z,\vec{\Omega};\tilde{\mathrm{V}}_\beta^0\right)=(4\pi)^{-1}\left(\kappa_{ext}\right)^3 P\left(\mathrm{V}_\beta^0\right)\delta\left(\vec{\tau}-\vec{\tau}^*\right)$ $\left(\vec{\tau}=\kappa_{ext}\vec{r},\ \vec{\tau}^*=\kappa_{ext}\vec{r}^*\right)$. Place the origin of a dimensionless rectangular right-hand Cartesian coordinate system $O\tilde{X}\tilde{Y}\tilde{Z}$ into the point where the source is (i.e. assume that $\vec{\tau}^*=\vec{0}$). Denote the optical distance from point O along a straight line to boundary S_β via $\tau_{\mathrm{S}_\beta}\left(\theta,\varphi\right)$ (the direction of this ray is set by angles θ,φ in the spherical coordinate system matching with system $O\tilde{X}\tilde{Y}\tilde{Z}$). At this, let the source be disposed in such a way that any ray originating from point O intersects S_β only once. For the above assumptions, one can prove the validity of the below formula (Rogovtsov, 1994) by using Eqs (7.86)–(7.88):

$$F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)=\frac{1-\omega_0}{4\pi}P\left(\mathrm{V}_\beta^0\right)\int\limits_0^{2\pi}d\varphi$$
$$\times\int\limits_0^{\pi}\sin\theta\left[\tau_{\mathrm{S}_\beta}\left(\theta,\varphi\right)\tilde{G}_{\infty;0}\left(\tau_{\mathrm{S}_\beta}\left(\theta,\varphi\right),\omega_0\right)+\int\limits_{\tau_{\mathrm{S}_\beta}(\theta,\varphi)}^{+\infty}\tilde{G}_{\infty;0}\left(\tau',\omega_0\right)\,d\tau'\right]d\theta. \tag{7.90}$$

Quantity $F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)$ can be calculated for any phase function by means of Eqs (7.88) and (7.90). Equation (7.90) is convenient to find the asymptotic formulae for quantity$F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)$. Denote the least upper bound of the set of the values of function$\tau_{\mathrm{S}_\beta}\left(\theta,\varphi\right)$, when $\theta\in[0,\pi]$ and $\varphi\in[0,2\pi]$, via $\tau_1\left(\mathrm{S}_\beta\right)$. One can get the following asymptotic formula from Eq. (7.90) (Rogovtsov, 1994, 1997):

$$F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)=\frac{1-\omega_0}{4\pi}P\left(\mathrm{V}_\beta^0\right)$$
$$\times\left[\sum_{l=1}^{r}c_l\left(\omega_0;0\right)\int\limits_0^{2\pi}d\varphi\int\limits_0^{\pi}\left(\tau_{\mathrm{S}_\beta}\left(\theta,\varphi\right)+\left(k_l\left(0\right)\right)^{-1}\right)\exp\left(-k_l\left(0\right)\tau_{\mathrm{S}_\beta}\left(\theta,\varphi\right)\right)\sin\theta\,d\theta\right]$$
$$+O\left(\Delta\right),\ \tau_1\left(\mathrm{S}_\beta\right)\to+\infty\text{ or }\omega_0\to1. \tag{7.91}$$

Here r is the number of the found roots of equation $\wp_0\left(-k^2;0\right)=0$ (this equation has at least one root); $c_l\left(\omega_0;0\right)=\left[\int_{-1}^{1}\mu\left(\Phi_0\left(-ik_l\left(0\right),\mu\right)\right)^2\,d\mu\right]^{-1}$ is the normalization constant corresponding to root $k_l\left(0\right)\in(0,1)$ (eigenfunctions $\Phi_0\left(-ik_l\left(0\right),\mu\right)$ of the reduced characteristic equation of the RTT are normalized by the condition of $\int_{-1}^{1}\Phi_0\left(-ik_l\left(0\right),\mu\right)\,d\mu=1$); $\Delta=\left(1-\omega_0\right)P\left(\mathrm{V}_\beta^0\right)\int_0^{2\pi}d\varphi\int_0^{\pi}\left(\tau_{\mathrm{S}_\beta}\left(\theta,\varphi\right)+\left(k^*\right)^{-1}\right)\times\exp\left(-k^*\tau_{\mathrm{S}_\beta}\left(\theta,\varphi\right)\right)\sin\theta\,d\theta$, where k^* is any number belonging to $\left(k_r(0),1\right)$, if there exist only r roots; and $k^*=k_{r+1}\left(0\right)$, if there exist more than r roots. The asymptotic formula of Eq. (7.91) is substantially simplified, if surface S_β has constructive symmetry with respect to the position of the point isotropic source.

If S_β is a sphere with optical radius τ_0 and the source is disposed in the symmetry center, then Eq. (7.91) assumes the form:

$$F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)=\left(1-\omega_0\right)P\left(\mathrm{V}_\beta^0\right)\sum_{l=1}^{r}c_l\left(\omega_0;0\right)\left(\tau_0+\left(k_l\left(0\right)\right)^{-1}\right)\times\exp\left(-k_l\left(0\right)\tau_0\right)+O\left(\Delta\right),\tag{7.92}$$

where $\tau_0\to+\infty$ *or* $\omega_0\to1$, $\Delta=\left(1-\omega_0\right)P\left(\mathrm{V}_\beta^0\right)\left(\tau_0+\left(k^*\right)^{-1}\right)\exp\left(-k^*\tau_0\right)$.

Let body V_β be a spheroid with optical lengths of its semi-axes being $\tilde{a},\tilde{a},\tilde{c}$ (boundary S_β of this body is set by equation $\tilde{x}^2\left(\tilde{a}\right)^{-2}+\tilde{y}^2\left(\tilde{a}\right)^{-2}+\tilde{z}^2\left(\tilde{c}\right)^{-2}=1$ in system $O\tilde{X}\tilde{Y}\tilde{Z}$). If the point isotropic source is disposed in the center of the spheroid, then one gets the following asymptotic formulae (Rogovtsov, 1994, 1997) from Eq. (7.91) by means of the generalized Laplace method (see, for example, Olver, 1974):

$$F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)\sim\left(1-\omega_0\right)c_1\left(\omega_0;0\right)\left(k_1\left(0\right)\left(1-\zeta^2\right)^{-1}\right)P\left(\mathrm{V}_\beta^0\right)\times\exp\left(-k_1\left(0\right)\tilde{c}\right),\quad k_1\left(0\right)\tilde{c}\to+\infty\ \left(\zeta\in(0,1)\right);\tag{7.93}$$

$$F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)\sim\left(1-\omega_0\right)c_1\left(\omega_0;0\right)\xi\left(\xi^2-1\right)^{-\frac{1}{2}}\left(\left(2k_1\left(0\right)\right)^{-1}\pi\tilde{a}\right)^{\frac{1}{2}}P\left(\mathrm{V}_\beta^0\right)\times\exp\left(-k_1\left(0\right)\tilde{a}\right),\quad k_1\left(0\right)\tilde{a}\to+\infty\ \left(\xi\in(1;+\infty)\right).\tag{7.94}$$

One should assume that $\tilde{c}=\zeta\tilde{a}$ in Eq. (7.93) (in this case, V_β is an 'oblate' spheroid). Respectively, quantity $\tilde{c}=\xi\tilde{a}$ in Eq. (7.94) (in this case V_β is a 'prolate' spheroid). Note the the asymptotic formulae of Eqs (7.93) and (7.94) are also valid when $\omega_0\to1$.

If one lets parameter ξ tend to infinity in Eq. (7.94), then the asymptotic formula for the case of an infinite circular cylinder with optical radius $\tilde{a}$ is obtained (the point isotropic source is disposed on its symmetry axis). This formula has the following form:

$$F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right)\sim\left(1-\omega_0\right)c_1\left(\omega_0;0\right)\left(\left(2k_1\left(0\right)\right)^{-1}\pi\tilde{a}\right)^{\frac{1}{2}}P\left(\mathrm{V}_\beta^0\right)\times\exp\left(-k_1\left(0\right)\tilde{a}\right),\ k_1\left(0\right)\tilde{a}\to+\infty.\tag{7.95}$$

Assume now that body V_β has a cubic shape, in the symmetry center of which there is disposed a point isotropic source. Then one can get the following asymptotic formula (Rogovtsov, 1994) from Eq. (7.91) by means of the generalized Laplace method and of a number of elementary transformations:

$$F_2\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right);\mathrm{V}_\infty\right) \sim 3\left(1-\omega_0\right)c_1\left(\omega_0;0\right)\left(k_1\left(0\right)\right)^{-1}P\left(\mathrm{V}_\beta^0\right) \times \exp\left(-2^{-1}k_1\left(0\right)\tilde{d}\right), k_1\left(0\right)\tilde{d}\to+\infty. \tag{7.96}$$

where $\tilde{d}$ is the optical length of the cube side.

As the conclusion of this subsection, let us give the simplest asymptotic formulae for monochromatic luminosity values of non-concave, optically thick, almost conservatively scattering, turbid bodies containing point isotropic sources at large optical distances from all points of their boundaries. For this, we will assume that the phase function is constant (it equals to unity to correspond to the isotropic scattering) or is forwardly extended. Taking into account Eqs (7.77), (7.91)–(7.95), the definition of normalization constant $c_1\left(\omega_0;0\right)$, and the asymptotic formula of function $\Phi_0\left(-ik_1\left(0\right),\mu\right)$ for $\omega_0\to 1$ (see, for example, Sobolev (1975) and Yanovitskij (1997)), one gets the following asymptotic formulae for the above-mentioned conditions (Rogovtsov, 1994, 1997):

$$F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \sim 2k_1\left(0\right)\tau_0 P\left(\mathrm{V}_\beta^0\right)\exp\left(-k_1\left(0\right)\tau_0\right), \upsilon_0\to 0,\ k_1\left(0\right)\tau_0\to+\infty. \tag{7.97}$$

(V_β has the shape of a sphere);

$$F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \sim 2\left(1-\zeta^2\right)^{-1}P\left(\mathrm{V}_\beta^0\right)\exp\left(-k_1\left(0\right)\tilde{c}\right), \upsilon_0\to 0,\ k_1\left(0\right)\tilde{c}\to+\infty. \tag{7.98}$$

(V_β has the shape of a oblate spheroid);

$$F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \sim \xi\left(\xi^2-1\right)^{-\frac{1}{2}}\left(2\pi k_1\left(0\right)\tilde{a}\right)^{\frac{1}{2}}P\left(\mathrm{V}_\beta^0\right)\exp\left(-k_1\left(0\right)\tilde{a}\right), \upsilon_0\to 0,\ k_1\left(0\right)\tilde{a}\to+\infty. \tag{7.99}$$

(V_β has the shape of an prolate spheroid);

$$F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \sim \left(2\pi k_1\left(0\right)\tilde{a}\right)^{\frac{1}{2}}P\left(\mathrm{V}_\beta^0\right)\exp\left(-k_1\left(0\right)\tilde{a}\right),\ \upsilon_0\to 0,\ k_1\left(0\right)\tilde{a}\to+\infty. \tag{7.100}$$

(V_β has the shape of an infinite circular cylinder);

$$F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \sim \left(2\pi\right)^{-1}k_1\left(0\right)P\left(\mathrm{V}_\beta^0\right)\int_0^{2\pi}d\varphi\int_0^{\pi}\tau_{\mathrm{S}_\beta}\left(\theta,\varphi\right) \times \exp\left(-k_1\left(0\right)\tau_{\mathrm{S}_\beta}\left(\theta,\varphi\right)\right)\sin\theta d\theta, \upsilon_0\to 0,\ k_1\left(0\right)\tau_0\left(\mathrm{S}_\beta\right)\to+\infty, \tag{7.101}$$

(V_β is a non-concave body).

Quantity $k_1\left(0\right)$ can be replaced by $\left(3\left(1-f_1\right)\left(1-\omega_0\right)\right)^{1/2}$ in all Eqs (7.97)–(7.101). These asymptotic expressions demonstrate the effect of the shape of a turbid medium on monochromatic luminosity $F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$ in the explicit and simple analytical form.

The asymptotic formulae of Eqs (7.97) to (7.101) describe the dependences of radiances of turbid bodies on dimensionless optical parameters defining the shape and optical sizes of these bodies. On the basis of these formulae, one can get the asymptotic expressions for mean radiative flux leaving turbid media via the portions of their boundaries with unit area. The said mean radiative flux is defined by the following relation $F_{mid}\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) = \left(Ar\left(\mathrm{S}_\beta\right)\right)^{-1} F_1\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$, where $A\,r\,\left(\mathrm{S}_\beta\right)$ is the area of boundary S_β of body V_β. Using this definition, explicit expressions for the areas of the boundaries of spheroids, and asymptotic formulae of Eqs (7.98) and (7.99), one derives the following asymptotic relations:

$$F_{mid}\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \sim \frac{2\left(\kappa_{ext}\right)^2 P\left(\mathrm{V}_\beta^0\right)\left(\tilde{c}\right)^{-2}\exp\left(-k_1(0)\tilde{c}\right)}{\pi\gamma_1^2\left[2\zeta^{-2}+\gamma_1^{-1}\ln\left(\left(1+\gamma_1\right)/\left(1-\gamma_1\right)\right)\right]},\ \upsilon_0 \to 0,\ k_1\left(0\right)\tilde{c} \to +\infty; \tag{7.102}$$

$$F_{mid}\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \sim \frac{\left(\kappa_{ext}\right)^2\left(k_1\left(0\right)\right)^{\frac{1}{2}} P\left(\mathrm{V}_\beta^0\right)\left(\tilde{a}\right)^{-\frac{3}{2}}\exp\left(-k_1(0)\tilde{a}\right)}{\sqrt{2\pi}\xi\left[\left(\pi/2\right)-\arccos\left(\xi^{-1}\gamma_2\right)+\xi^{-2}\gamma_2\right]},$$
$$\upsilon_0 \to 0,\ k_1\left(0\right)\tilde{a} \to +\infty; \tag{7.103}$$

where $\gamma_1 = \left(1-\zeta^2\right)^{1/2}$, $\gamma_2 = \left(\xi^2-1\right)^{1/2}$. Equation (7.102) corresponds to the case of an oblate spheroid, and Eq. (7.103) does the same for the case of a prolate spheroid.

If body V_β has a spherical shape, then the validity of the asymptotic formula follows from Eq. (7.97) and the definition of quantity $F_{mid}\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right)$:

$$F_{mid}\left(\mathrm{S}_\beta\left(\mathrm{V}_\beta^0\right)\right) \sim \left(2\pi\right)^{-1}\left(\kappa_{ext}\right)^2 P\left(\mathrm{V}_\beta^0\right) k_1\left(0\right)\tau_0^{-1}\exp\left(-k_1(0)\tau_0\right),$$
$$\upsilon_0 \to 0,\ k_1\left(0\right)\tau_0 \to +\infty.$$

Due to the spherical symmetry of the problem statement, this formula is also an asymptotic relation for the true radiative flux leaving any unit-area portion of the boundary of sphere V_β.

Note in addition that, via the right-hand parts of all the above-written asymptotic relations for mean radiation fluxes, one can easily find the lower-bounded estimates of the main terms of the asymptotics for mean radiation densities at the boundaries of turbid media having the shapes of oblate or prolate spheroids and sphere. To do so, it is enough to divide the said right-hand parts by the velocity of light in the turbid media indicated.

7.4.3 On the depth regimes of radiation fields and on the derivation of asymptotic expressions for mean emission durations of optically thick, turbid bodies

The depth regimes of radiation fields is intended to mean situations where radiation sources and regions with observation points are spaced from each other at large optical distances. For such situations, the simplification of functional dependences of radiation field characteristics on spatial, angular, temporal and other variables can occur due to a large number of mean scattering events, in which photons emitted by the sources and coming to an observation point participate. The

most substantial simplifications arise when the shape of a turbid medium, angular patterns, and radiation source dispositions possess by the same kind of symmetry. In this case, the total or partial separation of variables in the main terms of the asymptotic expressions for radiation field characteristics occurs. Such a separation is implemented, in particular, under the depth regimes for optically thick, turbid media having configurations with the plane-parallel, spherical or cylindrical symmetry (see, for example, the publications (Sobolev, 1975; Kolesov, 1983, 1985, 1986; Rogovtsov and Samson, 1985a; Yanovitskij, 1997; Kokhanovsky, 2006) and references therein). It should be noted that boundary-value problems of the RTT are essentially one-dimensional ones with respect to spatial variables, provided that the said symmetry is in place. Investigation of the depth regimes for multi-dimensional (over spatial variables) boundary-value problems of the RTT is a much more complicated problem as compared with the solutions to the similar problems for turbid media, having configurations with the above-mentioned kind of the symmetry. The study of the depth regimes, when the initial boundary-value problem is a multi-dimensional one, is interesting for testing various versions of numerical methods and the Monte Carlo procedure designed to solve problems of the RTT for turbid media with complicated configurations. On the other hand, while considering the depth regimes, it is often possible to derive analytical (semi-analytical) forms of the relationships for multiple light scattering in turbid media of various configurations. The asymptotic radiative transfer theory, which enables one to completely or partially describe the depth regimes for arbitrary phase functions and turbid media of complex configurations, has been elaborated in the papers (Rogovtsov, 1986c, 1988a, 1988c, 1988d, 1990b, 1991a, 1992a, 1994) on the basis of the GIRRM and in the papers (Germogenova and Pavelyeva, 1991; Pavelyeva,1990; Germogenova and Pavelyeva, 1994) by using functional analysis techniques. It has been proved in these publications that the investigation of the depth regimes of radiation fields for a number of initial three-dimensional (over spatial variables) boundary-value problems of the RTT can be reduced to the solution of the two- or one-dimensional boundary-value problems for the scalar radiative transfer equations that correspond to some model anisotropically absorbing turbid media having the shape of a plane-parallel slab or of the infinite cylindrical body. Moreover, in the papers (Rogovtsov, 1986c, 1988a, 1988d, 1990b), there have been detected situations, in which the main terms of the asymptotic formulae for radiation intensities can be expressed explicitly via elementary functions, the solution to the Milne problem (Sobolev, 1975; Yanovitskij, 1997), and the first eigenfunction of the reduced characteristic equation of the RTT that corresponds to the zeroth azimuthal harmonics. It should be especially emphasized that the normalization factors of all the main terms of the above-said asymptotic formulae for the depth regimes were derived by using the GIRRM only.

In publications (Rogovtsov, 1986c, 1988a, 1988c, 1988d, 1990a, 1990b, 1991a, 1991c, 1992a) it has been shown that on the basis of the GIRRM the depth regimes can be investigated for the following cases:

(i) V_β is a semi-infinite (in particular, two-layer) turbid medium, limited by a specularly reflecting surface and exposed by a point mono-directional source (see Fig. 7.7; in Fig. 7.7 τ_0 is optical depth; $\tau_0 \gg 1$).

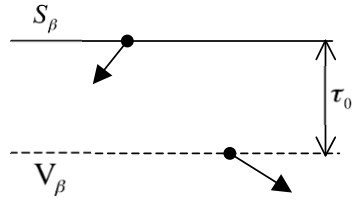


Fig. 7.7. Geometry of problems for case (i).

(ii) V_β is an optically thick, turbid layer exposed by a point mono-directional source (see Fig. 7.8; in Fig. 7.8 τ_0 is optical thickness layer; $\tau_0 \gg 1$).

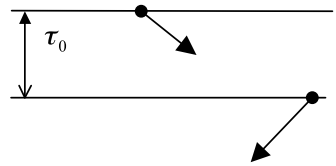


Fig. 7.8. Geometry of problems for case (ii).

(iii) V_β is a sphere containing a point mono-directional source at the center (see Fig. 7.9; τ_0 is the optical radius of this sphere; $\tau_0 \gg 1$).

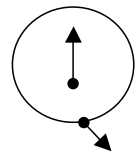


Fig. 7.9. Geometry of problems for case (iii).

(iv) V_β is a circular, infinite cylinder containing a linear mono-directional source at the symmetry axis (see Fig. 7.10; τ_0 is the optical radius of this cylinder; $\tau_0 \gg 1$).

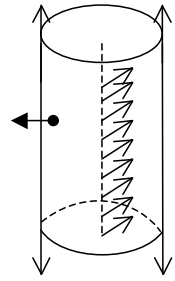


Fig. 7.10. Geometry of problems for case (iv)

(v) V_β is a cylindrical, turbid medium radiating a point mono-directional source (τ_0 is the minimal optical distance between this source and an observation point; $\tau_0 \gg 1$). In Fig. 7.11(a) V_β is an infinite cylindrical medium. In Fig. 7.11(b) V_β is a semi-infinite cylindrical medium. In Fig. 7.11(c) V_β is an optically thick, cylindrical medium.

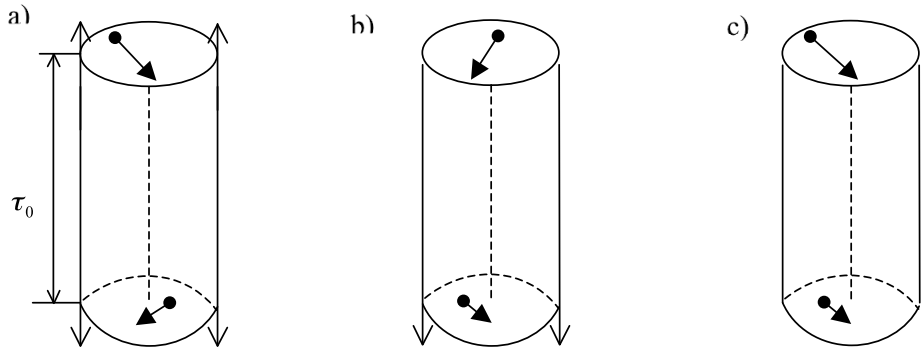


Fig. 7.11. Geometry of problems for case (v).

There have been derived a number of strict general expressions and double-sided inequalities for mean emission times t^* of turbid media in subsection 7.3.5 on the basis of original constructions of the GIRRM. At this, the said quantities have been expressed or estimated (from both the top and the bottom) via the solutions to the boundary-value problems for Eq. (7.10). Note now once more the important analytical result that occurs for the case of an arbitrary limited phase function and almost conservatively scattering, optically thick, macroscopically homogeneous turbid medium V_β bounded by a smooth surface S_β (let S_β be not an underlying surface) and containing non-stationary sources. For these conditions, taking into account the asymptotic formulae of Eq. (7.76), there follows from the inequalities of Eq. (7.62) the validity of the asymptotic expression: $t^* \sim U_1(U_0)^{-1}$, $\upsilon \to 0$, $k_1(0)\tau_0(S_\beta) \to +\infty$. In this expression, quantities U_0 and U_1 should be found by using the formulae of Eq. (7.55) and (7.56), in which $S_{\beta_2} = S_\beta$ and $V_{\beta_4} = V_\infty$ should be set. Note particularly that, although this expression formally resembles Eq. (7.59) (it is valid for conservative scattering only), quantities U_0 and U_1 differ from the relations of Eqs (7.60) and (7.61). Note also that the publications (Rogovtsov, 1988b, 1994, 2002) derived a number of simple asymptotic formulae and double inequalities for quantity t^* for the case of conservatively scattering, optically thick, turbid media having the shape of a spheroid (in particular, of a sphere). Besides, on the basis of the GIRRM, there have been obtained a number of generalizations of the flux integral and the K-integral (see, for example, Sobolev (1975) and Yanovitskij (1997)) for the case of turbid media of complicated configurations illuminated by non-stationary sources in Rogovtsov (1993). Anisimov and Rogovtsov (2002) use the result of Rogovtsov (1993) for solving an inverse problem on the determination of the absorption coefficient.

7.5 Conclusion

The main ideas and constructions of the GIRRM have been stated in this chapter. In addition using the GIRRM, a number of analytical, semi-analytical and numerical solutions to various boundary value problems of the RTT for turbid media of different configurations are here described and derived. It should be noted that the heuristicity and the efficiency of the GIRRM are due to several reasons. First, rather an abstract treatment of the GIP exceeded the limit of the classical considerations of the invariance principles, which were stated in theoretical physics, mathematics, and radiative transfer theory, was assumed as a basis of the GIRRM. This fact enables one to immediately treat the whole class of boundary-value problems of the RTT and to select them as a function of the conditions of the statements of specific RTT problems. Second, the GIP that is the basic element of the GIRRM permits simple enough physical interpretation. Third, the scheme of getting the GIRs within the scope of the GIRRM is rather flexible to be closely connected with the main considerations of ray optics and probabilistic interpretation of the RTT. Fourth, the GIRRM enables a certain hierarchy and relationship between the solutions to boundary-value problems of the RTT for turbid media of various configurations to be set forth. This fact provides an opportunity to reduce the solution to complicated boundary-value problems of the RTT to the consideration of simpler boundary-value problems of the RTT. Fifth, no assumptions on the symmetry and on the problem dimension with respect to spatial and other variables are initially laid down on the GIRRM constructions while the boundary-value problems are stated. This feature of the GIRRM has enabled a number of analytical and semi-analytical solutions to RTT problems to be obtained for turbid media, the configurations of which possess no substantial symmetry. The presence of symmetry of such a kind can just give additional opportunities for using the GIRRM. Sixthly, the GIRMM allows one to reformulate diverse complex direct and inversion problems of disperse medium optics and essentially to decrease the volume of mathematical calculations as compared with the volume needed to fulfill in direct solving the initial problems.

Results derived on the basis of the GIRRM can, in particular, be utilized for testing various numerical methods and Monte Carlo procedures. Besides, they permit one to qualitatively and quantitatively estimate the effect of the configuration of a turbid medium on radiation field characteristics therein. Moreover, having any meaningful information of the properties of the solutions to boundary-value problems of the RTT for a turbid medium of a certain configuration, it is possible, in principle, to get information on the similar or other solution properties of more complicated RTT problems by means of the GIRRM. In should also be mentioned that, by using the GIRRM, one can design a new version of the invariant embedding method suitable for solving various problems of mathematical physics (in particular, for solving multi-dimensional boundary-value problems of the RTT). A number of results in this direction have been obtained by Rogovtsov (1992b, 2008). It should be added to the above that RTT problems can be effectively solved by using the GIRRM for cases, when the boundaries of turbid media are underlying surfaces.

Acknowledgment

This chapter contains a number of results of scientific studies supported partially by the National Academy of Sciences of Belarus and by the Education Ministry of Belarus in the frame of the Fundamental Investigation Programs 'Wave' and 'Coherence'. The author would like to mention the positive impact had on the development of the GIRRM at different times by Professors V. A. Ambartsumian, M. A. El'yashevich, F. I. Fedorov, F. D. Gakhov, T. A. Germogenova, V. V. Ivanov, D. I. Nagirner, A. M. Samson, V. V. Sobolev, E. G. Yanovitskij, N. V. Konovalov and O. V. Pikichian during numerous scientific contacts and discussions with them.

The author would also like to thank Dr A. A. Kokhanovsky for constant interest in the results of the scientific studies and valuable discussions on the problems dealt with in this Chapter.

Appendix A: Main mathematical notations, conceptions, and constructions used while stating the general invariance principle and deriving the general invariance relations

A set is one of the major general mathematical notions. If A is a finite set, it will be written in the form of $\mathrm{A} = \{a_1,\ a_2,\ \ldots,\ a_n\}$, where $a_1,\ a_2,\ \ldots,\ a_n$ are its elements and n is the number of elements belonging to A. The belonging relation of an element b to a set B is denoted by a symbol $\in$ (i.e. $b \in B$). A general way to write any set X is the following presentation: $\mathrm{X} = \{x | C(x)\}$, where $C(x)$ is some condition (or their collections), to which the elements of this set should satisfy. If condition $C(x)$ does not hold (or $C(x)$ is a false assertion), we write $x \notin \mathrm{X}$ (i.e. x does not belong to X). In set theory, symbols $\subset$, $\cup$, $\cap$, $\setminus$ are used. Notation $\mathrm{A} \subset \mathrm{B}$ means that any element of the set A belongs to the set B (i.e. A is a subset of the set B). Sets A and B are considered to be equal, i.e. $\mathrm{A} = B$, if relations $\mathrm{A} \subset \mathrm{B}$, $\mathrm{B} \subset \mathrm{A}$ hold simultaneously. Symbol $\cup$ means the union operation of sets. It means that $\mathrm{A} \cup \mathrm{B} = \{x | x \in \mathrm{A} \text{ or } x \in \mathrm{B}\}$, i.e. set $\mathrm{A} \cup \mathrm{B}$ consists of all the elements belonging to set A, set B, or to the both these sets simultaneously. Expression $\mathrm{A} \cap \mathrm{B}$ defines an intersection of the sets, i.e. $\mathrm{A} \cap \mathrm{B} = \{x | x \in \mathrm{A} \text{ and } x \in \mathrm{B}\}$. Construction $\mathrm{A} \setminus \mathrm{B}$ means a difference of sets i.e. $\mathrm{A} \setminus \mathrm{B} = \{x | x \in \mathrm{A} \text{ and } x \notin \mathrm{B}\}$. An empty set, $\emptyset$, is thought to be a set containing no elements.

Another major general mathematical conception is a notion of map (function, operator) of set A into set B. Symbolically a map f is written as $f : A \to B$. Symbol f is thought to be a rule (law, algorithm), according to which a single, at least, element $b \in \mathrm{B}$ is specified for all $a \in D(f)$ $(D(f) \subset \mathrm{A})$. Symbol $D(f)$ is thought to be the domain of map f. For all $a \in D(f)$, the map f defines one of the elements of set B or their collection. If $b = f(a)$, where $a \in D(f)$ and $b \in \mathrm{B}$, then a is a prototype of element b and b is an image of element a. Image Imf of set A into set B is referred to as a set $Imf = \{b |\, b$ is the image of element a of the set $D(f)\}$. Set Imf is the range of map $f : \mathrm{A} \to \mathrm{B}$. The map $f : \mathrm{A} \to \mathrm{B}$ is referred to as a simple one, if for all $a \in D(f)$ there exists the unique image $b = f(a) \in \mathrm{B}$. This chapter considers simple maps only. Map $f : \mathrm{A} \to \mathrm{B}$ is said to be injective,

if for all $b \in Imf$ there exists only a single prototype $a \in D(f)$ $(b = f(a))$. The maps, utilized while stating the GIP, should not necessarily be injective ones. If map $f:\ \mathrm{A} \to \mathrm{B}$ is an injective one, $D(f) = \mathrm{A}$, and $Imf = \mathrm{B}$, a map is said to be a bijective map of set A onto set B. Map $f:\ \mathrm{A} \to \mathrm{B}$ generates an inverse map. By definition, the inverse map $f^{-1}:\ \mathrm{B} \to \mathrm{A}$ to the map $f:\ \mathrm{A} \to \mathrm{B}$ stands for such a map, in which an element $a \in \mathrm{A}$ is an image of an element $b \in \mathrm{B}$, iff $f(a) = b$ (in addition to $b \in Imf$). A map that is an inverse one to a simple map is not always a simple one. The maps utilized while stating the GIP are just such maps. But a map that is an inverse one to a bijective map is always a simple one. If relation $\mathrm{M} \subset D(f)$ is satisfied, then symbol $f(\mathrm{M})$ will be further assumed as set $f(\mathrm{M}) = \{x | x = f(a)\ \text{and}\ a \in \mathrm{M}\}$.

The map concept can be used for stating the sets themselves. Let B be some set, which one names as an indexing set. Assume that there is stated a simple map $f:\ \mathrm{B} \to V$, for which $D(f) = \mathrm{B}$ and $Imf = V$. In this case, for all $\beta \in \mathrm{B}$ there exists a single element $\mathrm{V}_\beta \in V$, for which $f(\beta) = \mathrm{V}_\beta$. At this, the equality of $V = \{\mathrm{V}_\beta | \beta \in \mathrm{B}\}$ is valid. If one considers any element $\mathrm{V}_\beta \in V$ as a set, then set V can be represented as a set coinciding with the union of all such elements, i.e. $V = \bigcup\limits_{\beta \in \mathrm{B}} \mathrm{V}_\beta$. Usually, this equality is written in the form of $V = \{\mathrm{V}_\beta\}_{\beta \in \mathrm{B}}$. Note that the simplest example of stating the sets by the above-described means is specifying numerical sequences $\{x_n\}_{n \in \mathrm{N}} = \{x_1,\ x_2,\ x_3, \ldots\}$, where N is the set of the natural numbers.

Once more important general mathematical concept is the notion of an algebraic operation. A unary algebraic operation specified on set A is meant as a simple map $f:\ A \to A$, for which $D(f) = A$. This operation maps any element a of set A to a single element $f(a)$ of the same set. Denote a set of all unary algebraic operations specified on set A via W_A. Let $f_1,\ f_2 \in \mathrm{W}_A$ (i.e. maps $f_1:\ A \to A$ and $f_2:\ A \to A$ are simple maps and $D(f_1) = D(f_2) = A$). Then a composition $f_3 = (f_2 f_1)$ of these two unary algebraic operations will be assumed as a map $f_3:\ A \to \mathrm{A}$ that assigns any element $a \in \mathrm{A}$ to a single element $c \in \mathrm{A}$ according to the rule of $c = f_3(a) = f_2(f_1(a))$, where $b = f_1(a) \in \mathrm{A}$. It follows from the definition of a unary (algebraic) operation that this composition is also a unary operation specified on set A. Fig. 7.12 simplifies graphically the sense of unary operations $f_2,\ f_1$ and their composition $(f_2 f_1)$.

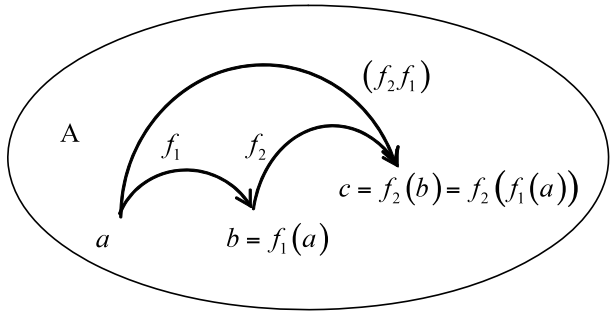


Fig. 7.12. Graphical illustration of operation f_2, f_1 and $(f_2 f_1)$.

It follows from the above-said that $(f_2 f_1) \in \mathrm{W}_A$. The composition operation of two maps belonging to set W_A represents an important example of a binary algebraic operation. State the general definition of a binary algebraic operation. Let X be a non-empty set. We will say that a binary algebraic operation $*$ is specified on set X, if any ordered pair (a, b), where $a, b \in \mathrm{X}$, is assigned to a single element $c \in \mathrm{X}$, i.e. the equality of $c = a * b$ is valid. An operation $*$ is called an associative one, if the equality of $(a*b)*c = a*(b*c)$ is true for all $a, b, c \in \mathrm{X}$. A binary algebraic operation $*$ is named as a commutative one, if the equality of $a*b = b*a$ is true for all $a, b \in \mathrm{X}$. Especially note that a binary algebraic operation of the composition of two unary operations is an associative one (see, for example, Birkhoff and Bartee, 1970). Due to this, the equality of $(f_{\alpha_1} f_{\alpha_2}) f_{\alpha_3} = f_{\alpha_1}(f_{\alpha_2} f_{\alpha_3})$ is valid for any $f_{\alpha_1}, f_{\alpha_2}, f_{\alpha_3} \in \mathrm{W}_A$. However, generally speaking, this composition operation is not a commutative one. Since any non-empty set, on which an associative binary algebraic operation is specified, is a semi-group by definition, set W_A introduced above is a semi-group too. It should be underlined that semi-group $\mathfrak{R}$ introduced in subsection 7.2.2 can be treated as a subsemi-group of semi-group W_A, when one takes set $\{\mathrm{V}_\beta\}_{\beta \in \mathrm{B}}$ as set A. At this, a subsemi-group of some semi-group stands for any of its subset that is a semi-group itself.

The conception of an invariant as well as symmetry (invariance) properties and principles play an important role in various scientific fields. Their descriptions and statements are usually made with using the notions of a group and a relation of equivalence. Set X is named as a group, if it is a semi-group and the following conditions are satisfied:

1. there exists a neutral element e in set X, i.e. the equalities of $e * a = a * e = a$ are true for all $a \in \mathrm{X}$;
2. for all $a \in \mathrm{X}$, there exists an inverse element a^{-1}, for which the equalities of $a^{-1} * a = a * a^{-1} = e$ are true.

Groups, the elements of which are bijective maps (transformations) of sets to themselves, occupy a special position among groups widely utilized while studying and defining symmetry (invariance) properties of various objects, constructions, geometrical figures, equation solutions, etc. In addition, various subgroups of such groups (any subset of a group, that is a group itself, stands for a subgroup of the original group) are used for the same purposes. Go to the definition of a relation of equivalence specified on some non-empty set X. The relation conception as well as the set, the map, and the algebraic operation relations is one of the most general and important mathematical notions. We will assume binary relation χ (or simply relation) as any non-empty subset of a set of ordered pairs (i.e. of set $\mathrm{X} \times \mathrm{X} = \{(\mathrm{a}, \mathrm{b}) | a \in \mathrm{X} \text{ and } b \in \mathrm{Y}\}$). Let expressions $(a, b) \in \chi$ and $a\chi b$ be equivalent (interchangeable). The examples of the relations written down in the form of $a\chi b$ are $a = b$, $a < b$, $a > b$. Relation χ is called reflexive one, if there occurs $a\chi a$ for any $a \in \mathrm{X}$. Relation χ is called symmetric, if it follows $b\chi a$ from $a\chi b$ for any $a, b \in \mathrm{X}$. Relation χ is called transitive, if it follows $a\chi c$ from $a\chi b$ and $b\chi c$ for any $a, b, c \in \mathrm{X}$. Relation θ specified on non-empty set X will be called a relation of equivalence if it is reflexive, symmetric and transitive. Relation θ of equivalence partitions set X by non-intersecting subsets (classes of equivalence). These classes in pairs do not have common elements. Each of these classes of equivalence con-

sists of all elements of set X that are related by the relation θ of equivalence with some fixed element of the given class. The exclusive importance of the relation-of-equivalence conception is due to its specifying the generalized form of the equality of all objects (elements) from the same class of equivalence. This means that, in some sense, elements of the same class of equivalence possess the same properties. An invariant is often assumed in mathematics as a map $f : \mathrm{X} \to \mathrm{A}$, where X is a non-empty set with relation θ of equivalence specified on it. At this, all elements of any selected class of equivalence of set X should be mapped into a single element of set A. Different elements of set A should, however, correspond to different classes of equivalence of set X. Sometimes, while treating applied problems, an invariant is intended as the whole population of elements of A, into which set X with a relation of equivalence specified on it is mapped. It should be pointed out that the study of invariants is connected with classification problems for objects of one or another type. Notice that, in various scientific fields, one can provide the invariance (unchanging in some sense) conception with a sense that does not always coincide with the above-given classical definition of the invariance. It has been noted earlier that the conceptions of a group, of a relation of equivalence, and of symmetry (invariance) principles are connected with each other. This should be added with that group operations of groups (and of their subgroups) are often utilized as one of general and constructive elements, while specifying and stating relations of equivalences on sets. This is connected, in particular, with the fact that there is a bijective correspondence between the two conception sets of {reflexivity; symmetry; transitivity} and {identical transformation; direct and inverse transformations; composition of transformations}. The notions belonging to the first set are inherent to the relation of equivalence and the notions of the second set are used while defining groups of bijective maps. Symmetry (invariance) principles in mathematics and physics stand usually for statements pointing to the invariance (unchanging in some sense) of mathematical or physical objects, constructions, etc. with respect to one or another transformation groups.

As has been noted and proved by Rogovtsov (1981a, 1981b, 1989, 1994, 1999), not all the invariance principles can be regarded as symmetry principles. Generally, the invariance principles should be understood as statements postulating the invariance of sets of objects, constructions, solutions, properties, relations, etc. with respect to operations (actions) belonging to some set W (this set can generally be not a group and even a semi-group). Just such a general treatment was used, while stating the GIP, deriving the GIRs, and constructing the GIRRM. It should be added to the above that operations (actions) from set W can generally lead to some partial losses in the properties of original objects, constructions, solutions, relations, etc. Besides, operations (actions) from W can partially change one or another characteristic of the above-mentioned original sets. A situation of such a kind is implemented, in particular, under the action of operations belonging to semigroup $\mathfrak{R}$ on set $\{\mathrm{V}_\beta\}_{\beta\in\mathrm{B}}$ (see the statement of the GIP in subsection 7.2.2). Once more two general mathematical conceptions were utilized, while stating the GIP. These conceptions are 'bounded subset' and 'almost everywhere'. Term 'bounded subset' in subsection 7.2.2 is to be understood as any set in the three-dimensional Euclidean point space E^3 (see, for example, Faure et al. (1964)) that can be completely included in a sphere with a finite radius. The notion 'almost everywhere'

is closely related with the conceptions of a measure and of a measurable set (see, for example Kolmogorov and Fomin (1968) and Richtmyer (1978)). The 'measure' notion is the natural generalization of the concepts of a length of line segment, a plane figure area, a spatial figure volume, an increment of a non-decreasing function on a semi-open interval, etc. A set is called a measurable (according to Lebesgue), if one can uniquely, according to a certain rule, assign a measure to it (the strict definition of a measure is given, for example, by Kolmogorov and Fomin (1968)). A measure is a nonnegative function defined on a set of measurable sets and possessing the additivity property (i.e. a measure of a union of non-intersecting measurable sets is equal to the sum of the measures of each one of the sets). We will say that some property is fulfilled almost everywhere on set X, if the property is valid for all elements $x \in \mathrm{X}$, except for, maybe, elements belonging to some subset A of set X, the measure of this subset being zero. A measure in the three-dimensional Euclidean point space E^3 was used, while stating the GIP (see subsection 7.2.2). Within the frame of this measure, curves, surfaces, and countable sequences of points have the measure equal to zero. It should be clarified that the operations of kinds (ii) and (iii) introduced in subsection 7.2.2 do not make radiation fields invariant (unchangeable) on sets of zero measure in E^3.

References

Adams, M.L., Larsen, E.W., 2002: Fast iterative methods for discrete-ordinates particle transport calculation, *Prog. Nucl. Energy*, **40** (1), 3–159.

Adomian, G., Pandolfi, M., and Rach, R., 1988: An application of the decomposition method to the matrix Riccati equation in a neutron transport process, *J. of Math. Anal. and Appl.*, **136**, 557–567.

Ambartsumian, V.A., 1943a: On the question of diffuse light reflection from turbid medium, *Dokl. Akad. Nauk SSSR*, **38**, 257–261 (in Russian).

Ambartsumian, V.A., 1943b: On the problem of diffuse light reflection, *Zhurnal experimental'noj i teoreticheskoj fiziki*, issue 9–10, 323–334 (in Russian).

Ambartsumian, V.A., 1944a: Light diffusion through scattering medium with large optical thickness, *Dokl. Akad. Nauk SSSR*, **43**, 106–110 (in Russian).

Ambartsumian, V.A., 1944b: On the one-dimensional case of the problem on scattering and absorbing medium with finite optical thickness, *Izv. Akad. Nauk ArmSSR*, estestvenniye nauki, 1–2, 31–36 (in Russian).

Ambartsumian, V.A., 1947: On diffuse light reflection an transmission of anisotropic one-dimentional scattering medium with finite optical thickness, *Dokl. Akad. Nauk ArmSSR*, **7**, 199–202 (in Russian).

Ambartsumian, V.A., 1948: Turbid medium with homogeneous distribution of sources, *Dokl. Akad. Nauk ArmSSR*, **8**, 149–151 (in Russian).

Angel, E., Bellman, R., 1972: *Dynamic Programming and Partial Differential Equations*, New York: Academic Press.

Anisimov, V.Ya., Rogovtsov, N.N., 2002: Determination of the coefficient of light absorption via transformation of a pulse in a disperse medium, *J. Applied Spectroscopy*, **69**, 884–895.

Avaste, O.A., Viik, T.F. and Dmitrieva-Arago, L.R., 1978: On the absorption of solar, diffuse and reflected radiation in the visible region on the spectrum, *Izvestiya AN USSR. Fiz. Atmosf. i okeana*, **14**, 932–938 (in Russian).

Beauwens, R., Devooght, J., 1968: The study of one-speed multiregion transport problems in plane geometry by the method of boundary sources, *Nucl. Sci. and Eng.*, **32**, 249–261.

Bekefi, G., 1966: *Radiation Processes in Plasmas*, New York: John Wiley and Sons, Inc.

Bellman, R., 1986: A collection of works of R.E. Bellman (1920–1984), (Roth, R.S., ed.), New York: World Science.

Bellman, R., 1954: The theory of dynamic programming, *Bull. Am. Math. Soc.*, **60**, 503–516.

Bellman, R., Kalaba, R., 1956: On the principle of invariant imbedding and propagation through inhomogeneous media, *Proceedings of the National Academy of Sciences*, **42**, 629–632.

Birkhoff G., Bartee T.C., 1970: *Modern Applied Algebra*, New York: McGraw-Hill.

Bogolyubov, N.N., Shirkov, D.V., 1976: *Introduction to the Theory of Quantized Fields*, Moscow: Nauka (in Russian).

Case, K.M., 1957: Transfer problems and reciprocity principle, *Rev. Mod. Phys.*, **29**, 651–663.

Case, K.M., 1969: On the boundary value problems of linear transport theory, *SIAM. Proc. Amer. Math. Soc.*, **1**, 17–36.

Case, K.M. and Zweifel, P.F., 1967: *Linear Transport Theory*, Reading, MA: Addison-Wesley.

Case, K., de Hoffman, F., and Placzek, G., 1953: *Introduction to the Theory of Neutron Diffusion*, Vol. 1, Washington: U.S. Government Printing Office.

Cassell, J.S., Williams, M.M.R., 2006: Radiation transport and internal reflection in a sphere, *J.Q.S.R.T.*, **101**, 16–28.

Cassell, J.S., Williams, M.M.R., 2007: Radiation transport and internal reflection in a two region, turbid sphere, J.Q.S.R.T., **104**, 400–427.

Chandrasekhar, S., 1950: *Radiative Transfer*, London: Oxford University Press.

Davison, B., 1958: *Neutron Transport Theory*, London: Oxford University Press.

Dolginov, A.Z., Gnedin, Yu.N. and Silant'ev, N.A., 1995: *Propagation and Polarization of Radiation in Cosmic Media*, Amsterdam: Gordon & Breach.

Domke, H., 1976: Reduction of radiative transfer problems in semi-infinite media to linear Fredholm integral equations, *J.Q.S.R.T.*, **16**, 973–982.

Domke, H., 1978a: Linear Fredholm integral equations for radiative transfer problems in finite plane parallel media. I. Imbedding in an infinite medium, *Astron. Nachr., Bd.*, **299**, 87–93.

Domke, H., 1978b: Linear Fredholm integral equations for radiative transfer problems in finite plane parallel media. II. Imbedding in a semi-infinite medium, *Astron. Nachr., Bd.*, **299**, 95–102.

Domke, H., Yanovitskij, E.G., 1981: A simple computational method for internal polarized radiation fields of finite slab atmospheres, *J.Q.S.R.T.*, **26**, 389–396.

Edwards R.E., 1965: *Functional Analysis. Theory and Applications*, New York: Holt, Rinehart & Winston.

Einstein, A., 1965: Collected scientific works [Russian translation], vol. 1, Moscow: Nauka.

Einstein, A., 1966: Collected scientific works [Russian translation], vol. 2, Moscow: Nauka.

Ershov, Yu., P., Shikhov, S.B., 1972: Application of the integral transformations method to the solution of boundary-value problems of neutron transport. *Zhurnal vichislitel'noj matematiki i matematicheskoj fiziki*, **12**, 639–655 (in Russian).

Faure, F., Kaufmann, A., and Denis-Papin, M., 1964: *Mathematiques Nouvelles*, vol. 1, Paris: Dunod.

Fraenkel, A., Bar-Hillel, Y., 1958: *Foundations of Set Theory*, Amsterdam: North-Holland Publishing Company.

Freimanis, J., 2005: On Green's function for spherically symmetric problems of transfer of polarized radiation, *J.Q.S.R.T.*, **96**, 451–472.

Ganapol, B.D., 1976: Time-dependent moments of the monoenergetic transport equation in spherical and cylindrical geometry, *Nucl. Sci. Eng.*, **60**, 103–105.

Ganapol, B.D., Myneni, R.B., 1992: The application of the principles of invariance to the radiative transfer equation in a plant canopies, *J.Q.S.R.T.*, **48**, 321–339.

Ganapol, B.D., Kornreich, D.E., Dahl, D.A., Nigg, D.W., Jahshan, S.N., and Temple, C.A., 1994: The searchlight problem for neutrons in a semi-infinite medium, *Nucl. Sci. Eng.*, **118**, 38–53.

Ganapol, B.D., Kornreich, D.E., 1995: A 3-D neutron transport benchmark solution, *Transport Theory and Stat. Phys.*, **24**, 89–111

Ganapol, B.D., Pomraning, G., 1996: The two-region Milne problem, *Nucl. Sci. Eng.*, **123**, 110–120

Ganapol, B.D., McCormick, N.J., 1997: Radiative transfer half-space problems by the source function method, *Transport Theory and Stat. Phys.*, **26**, 507–532.

Ganapol, B.D., 2003: Fourier transform solution in spherical geometry, *Transport Theory and Stat. Phys.*, **32**, 587–605.

Ganapol, B.D., Kornreich, D.E., 2005: Multigroup 1D homogeneous medium benchmark using the Green's funsion method, Paper No. 61, 1–26, ANS/MEC, topical Meeting, Avignon, France (2005).

Ganapol, B.D., 2008: *Analytical Benchmarks for Nuclear Engineering Application*, Paris: OECD/NEA.

Garcia, R.D.M., Siewert, C.E., 1983: Multislab multigroup transport theory with Lth order anisotropic scattering, *J. Comput. Phys.*, **50**, 181–192.

Garcia R.D.M., Siewert, C.E., and Yacout, A.M., 2008: On the use of Fresnel boundary and interface radiative-transfer calculations for multilayerd media, *J.Q.S.R.T.*, **109**, 752–769.

Germogenova, T.A., 1986: *Local Properties of Solutions of the Transfer Equation*, Moscow: Nauka (in Russian).

Germogenova, T.A., Pavelyeva, E.B., 1991: The characteristic equation in radiative transfer problems in elongated cylindrical regions, *USSR Computational Mathematics and Mathematical Physics*, **29**, 158–170.

Germogenova, T.A., Pavelyeva, E.B., 1994: Semi-analytical benchmark calculations for two-and three-dimensional deep penetration problems, Proceeding of the 8th International Conference on Radiation Shielding, Arlington, Texas, USA, 24–28 April, vol. 1, 185–192.

Germogenova, T.A., Shulaya, D.A., 1976: Characteristic equation of the theory of radiation transport, *Soviet Physics Doklady*, **21**, 711–713.

Gibson, A.P., Hebden, J.C., and Arridge, S.R., 2005: Recent advances in diffuse optical imaging, *Phys. Med. Biol.*, 50, R1–R43.

Gorodnichev, E.E., Kuzovlev, A.I. and Rogozkin, D.B., 2006: Multiple scattering of polarized light in turbid media with large particles, in *Light Scattering Reviews*, vol. 1 (Kokhanovsky, A.A., ed.), 291–337, Berlin: Springer.

Hansen, J.E., 1969: Radiative transfer by doubling very thin layers, *Astrophys. J.*, **155**, 565–573.

Hovenier, J.W., van der Mee, C., and Domke, H., 2004: *Transfer of Polarized Light in Planetary Atmospheres. Basic Concepts and Practical Methods*, Dordrecht: Kluwer Academic Publishers.

Hunt, G.E., Grant, I.P., 1969a: Discrete space theory of radiative transfer I. Fundamentals, *Proc. Roy. Soc. Lond. A.*, **313**, 183–197.

Hunt, G.E., Grant, I.P, 1969b: Discrete space theory of radiative transfer and its application to problems in planetary atmospheres, *J. Atmos. Sci.*, **26**, 963–972.

Ibragimov, N.Kh., 1985: *Transformation Groups Applied to Mathematical Physics*, Dordrecht: D. Reidel.

Irvine, W.M., 1968: Diffuse reflection and transmission by cloud and dust layers, *J.Q.S.R.T.*, **8**, 471–185.

Ivanov, V.V., 1975: Invariance principles and internal radiation fields in semi-infinite atmospheres, *Astron. zhurn.*, **52**, 217–226 (in Russian).

Ivanov, V.V., 1976: Radiative transfer in infinite atmospheres, *Astrofizika*, **12**, 255–271; **12**, 369–379 (in Russian).

Ivanov, V.V., Volkov, E.V., 1978: Radiative transfer: invariance principles for the Green function, *Trudy Astron. Observ. LGU*, **35**, 3–30 (in Russian).

Jones W.B., Thron, W.J., 1980: *Continued Fractions. Analytic Theory and Applications*, London: Addison-Wesley.

Kadomtsev, B.B., 1957: Invariance principle for homogeneous medium of arbitrary geometrical shape, *Dokl. Akad. Nauk SSSR*, **112**, 831–834.

Kascas, A., Tezcan, C., Gülcyüz, M.C., 2000: The solution of the third form transport equation using singular eigen-functions: the slab and the sphere critically problems, *J.Q.S.R.T.*, **66**, 519–528.

Katsev, I.L., 1969: On integral characteristics for non-stationary light scattering, *Dokl. Akad. Nauk BSSR*, **13**, 118–121 (in Russian).

Klein, F., 1926: *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, vol. 1, Berlin: Verlag von Julius Springer.

Klose, A.D., 2009: Radiative transfer of luminescence light in biological tissue, in *Light Scattering Reviews* (A.A. Kokhanovsky, ed.), vol. 4, 293–345, Chichester, UK: Springer–Praxis.

Klyatskin, V.I., 1986: *Imbedding Method in Wave Propagation Theory*, Moscow: Nauka (in Russian).

Kokhanovsky, A.A., 2004: *Light Scattering Media Optics*, Berlin: Springer-Verlag.

Kokhanovsky, A.A., 2006: Asymptotic radiative transfer, in *Light Scattering Reviews*, vol. 1 (A.A. Kokhanovsky, ed.), 253–289, Berlin: Springer.

Kolesov, A.K., 1983: A point source in an absorbing and anisotropically scattering infinite homogeneous medium, *Dokl. Akad. Nauk SSSR*, **272**, 53–56 (in Russian).

Kolesov, A.K., 1985: Asymptotic relations in the theory of radiative transfer in a sphere and a spherical shell, *Astrophysics*, **22**, 107–114.

Kolesov, A.K., 1986: On radiative transfer in media with cylindrical symmetry, *Dokl. Akad. Nauk SSSR*, **278**, 115–118 (in Russian).

Kolmogorov, A.N., Fomin S.V., 1968: *Elements of Function Theory and Functional Analysis*, Moscow: Nauka.

Kovalev, V.F., Shirkov, D.V., 2008: Renormalization-group symmetries for solutions of nonlinear boundary-value problems, *Phys. Usp.*, **51**, 815–830.

Kuratowski, K., Mostowski, A., 1967: *Set Theory*, Amsterdam: North-Holland Publishing Company.

Laletin, N.T., 1969: Method of surface pseudosources for solution of neutron transfer equation, in *Computational Methods in Transfer Theory* (G.I. Marchuk, ed.), 228–245, Moscow: Atomizdat (in Russian).

Larsen, E.W., Pomraning, G.C., and Badham, V.C., 1980: On the singular eigenfuncions for linear transport in a exponential atmosphere, *J. Math. Phys.*, **21**, 2448–2454.

Latishev, V.N., 1984: Construction of asymptotic solution to the neutron transfer equation for cylindrical region in case of small free path length, *Differentsial'niye uravneniya*, **20**, 1596–1600, (in Russian).

Leroy, J.L., 2000: *Polarisation of Light and Astronomical Observations*, Amsterdam: Gordon and Breach Publs.

Liou, K.N., 2002: *An Introduction to Atmospheric Radiation*, New York: Academic Press.

Loyalka, S.K., 1974: A solution of the Milne problem using full-range completeness of Case's eingenfunctions, *Nucl. Sci. Eng.*, **54**, 353–356.

Marchuk, G.I., 1961: *Methods of Calculations of Nuclear Reactors*, Moscow: Atomizdat (in Russian).

Marchuk, G.I., Lebedev, V.I., 1971: *Numerical Methods in Neutron Transfer Theory*, Moscow: Atomizdat (in Russian).

Marchuk, G.I., et al., 1980: *The Monte Carlo Methods in Atmospheric Optics*, Berlin: Springer.

Marshak, A., Davis, A., 2005: *Three-Dimensional Radiative Transfer in Cloudy Atmospheres*, Berlin: Springer.

Maslennikov, M.V., 1968: Anisotropic scattering Milne's problem, *Trudy MIAN SSSR*, **97**, (in Russian).

McCormick, N.J., Sanchez R., 1981: Inverse problem transport calculation for anisotropic scattering coefficients, *J. Math. Phys.*, **22**, 199–208.

Mihalas, D., 1978: *Stellar Atmospheres* (2nd edn), San Francisco: Freeman.

Mikhailov, G.A., 1987: *Optimization of Weighting Monte Carlo Methods*, Moscow: Nauka (in Russian).

Minin, I.N., 1974: Light absorption in the atmosphere and on the planetary surface with multiple scattering and reflection, *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana*, **10**, 228–235 (in Russian).

Minin, I.N., 1988: *Theory of Radiative Transfer in Planetary Atmospheres*, Moscow: Nauka (in Russian).

Mnatsakanian, M.A., 1982: Non-linear transfer problems and renormalization group, *Dokl. Akad. Nauk SSSR*, **262**, 856–860 (in Russian).

Morel, J.E., McGhee, J.M., and Larsen E.W., 1996: A three-dimensional time-dependent unstructured tetrahedral-mash SP_N method, *Nucl. Sci. Eng.*, **123**, 319–327.

Nagirner, D.I., 1986: On radiative transfer in a layer and sphere and a cylinder, *Dokl. Akad. Nauk SSSR*, **276**, 606–609 (in Russian).

Nagirner, D.I., 1994a: Radiative transport in a cylinder. I. Resolvent of the fundamental integral equation, *Astrophysics*, **37**, 69–78.

Nagirner, D.I., 1994b: Radiative transfer in a cylinder. II. Special problems. Asymptotics, *Astrophysics*, **37**, 362–371.

Nagirner, D.I., 2001: *Lections on Radiative Transfer Theory*, St-Petersburg: St-Petersburg University Publishing Company (in Russian).

Nikoghossian, A.G., Haruthyunian, H.A., 1989: The application of the invariance principle to the solution of various problems of non-coherent and non-isotropic scattering, in *Invariance Principle and its Applications* (M.A. Mnatsakanian and O.V. Pikichian, eds), 232–241, Erevan: Izd. Akad. Nauk ArmSSSR (in Russian).

Nikoghossian, A.G., 2009: On the 100th anniversary of the birth of Academician V.A. Ambartsumian, *Astrophysics*, **52**, 1–23.

Nikolaeva, O.V. et al., 2007: Radiative transfer in horizontally and vertically inhomogeneous turbid media, in *Light Scattering Reviews*, vol. 2, (A.A. Kokhanovsky, ed.), 295–347, Berlin: Springer.

Olver, F.W.J., 1974: *Asymptotics and Special Functions*, New York: Academic Press.

Ovsyannikov, I.V., 1982: *Group Analysis of Differential Equations*, New York: Academic Press.

Pavelyeva, E.B., 1990: Asymptotics of solution to the equation of radiation transport in a disk of large radius with source near the symmetry exist, Preprint No. 97, Moscow: IPM Akad. Nauk SSSR (in Russian).

Pikichian, O.V., 1982a: The general invariance relations to study transfer problems in a media of geometrical and physical characteristics of arbitrary complexity, *Dokl. Akad. Nauk SSSR*, **262**, 860–863 (in Russian).

Pikichian, O.V., 1982b: Determing the radiation field in arbitrarily chosen macroscopic parts of component medium using known solutions to transfer problems, *Dokl. Akad. Nauk SSSR*, **263**, 601–606 (in Russian).

Pikichian, O.V., 1983: Finding the solution of the problem of transfer in isolated part of volume using known solutions for total volume, *Dokl. Akad. Nauk SSSR*, **273**, 861–865 (in Russian).

Pikichian, O.V., 1989: Principles of invariance and transfer problems with physical and geometrical characteristics of arbitrary complexity, in *Invariance Principle and its Applications*, (M.A. Mnatsakanian and O.V. Pikichian, eds), 88–119, Erevan: Izd. Akad. Nauk ArmSSR (in Russian).

Redheffer, R.M., 1954: Novel uses of functional equations, *J. Rat. Mech. Anal.*, **3**, 271–279.

Richtmyer, R.D., 1978: *Principles of Advanced Mathematical Physics*, vol. 1, New York, Springer-Verlag.

Riesz, F., Sz.-Nagy, B., 1972: *Lecons d'Analyse Fonctionelle*, Budapest: Akademiai Kiado.

Rogovtsov, N.N., 1980: Reconstruction of the internal radiation field distribution from its characteristics at the scattering medium boundaries, *Izvestiya, Atmospheric and Oceanic Physics*, **16**, 160–165.

Rogovtsov, N.N., 1981a: Partial invariant determination and the relation between the radiation fields within a scattering volume and at the boundaries, *J. Applied Spectroscopy*, **34**, 335–342.

Rogovtsov, N.N., 1981b: Transfer theory and the general invariance principle, *Dokl. Akad. Nauk BSSR*, **25**, 420–423 (in Russian).

Rogovtsov, N.N., 1981c: Some applications of the general principle of invariance in the theory of radiation transfer, *J. Applied Spectroscopy*, **35**, 1044–1050.

Rogovtsov, N.N., 1983: Invariance relations and radiative transfer problems for complex configurations media, *Dokl. Akad. Nauk BSSR*, **27**, 34–37 (in Russian).

Rogovtsov, N.N., 1985a: The application of general invariance relations to the theory of radiative transfer for complex shaped media, in *Abstract Book 3rd International Symposium on Radiation Physics*, p. 75, September 30 – October 4, Ferrara (Italy).

Rogovtsov, N.N., 1985b: On a calculation of the characteristics of radiation fields in scattering objects of complicated shapes on the basis of general invariance relations, *Izvestiya, Atmospheric and Oceanic Physics*, **21**, 846–847.

Rogovtsov, N.N., 1985c: Radiation transfer in canonical absorbing and scattering media, *J. Applied Spectroscopy*, **42**, 581–584.

Rogovtsov, N.N., 1985d: A general way of solving for radiation transport for objects of any configuration, *J. Applied Spectroscopy*, **43**, 813–816.

Rogovtsov, N.N., 1986a: Radiative transfer in an optically thick sphere, *J. Applied Spectroscopy*, **44**, 415–418.

Rogovtsov, N.N., 1986b: On calculation of radiative fluxes in scattering bodies of a non-concave form, *Dokl. Akad. Nauk. BSSR*, **30**, 609–613 (in Russian).

Rogovtsov, N.N., 1986c: Asymptotical formulae for radiation fields in spherically and cylindrically symmetrical optically thick scattering media. *Dokl. Akad. Nauk BSSR*, **30**, 901–904 (in Russian).

Rogovtsov, N.N., 1988a: On asymptotic expression for radiation fields in semi-infinite and optically thick scattering media, *Dokl. Akad. Nauk BSSR*, **32**, 413–416 (in Russian).

Rogovtsov, N.N., 1988b: To the determination of mean emission time of optically thick sphere containing arbitrary internal sources, *Vesti Akad. Navuk. BSSR, issue fiz.-matem. navuk*, No. 2, 101–106 (in Russian).

Rogovtsov, N.N., 1988c: An asymptotic expression for the surface Green function of the radiative transfer equation for the case of a homogeneous semi-infinite medium, *Dokl. Akad. Nauk BSSR*, **32**, 1085–1088 (in Russian).

Rogovtsov, N.N., 1988d: Some asymptotic expressions for the radiation fields in plane homogeneous media containing point sources, *Astrophysics*, **29**, 781–787.

Rogovtsov, N.N., 1989: Invariance principle applications for media of arbitrary configuration: Algebraic treatment, Applications, in *Invariance Principle and its Applications* (M.A. Mnatsakaniyan and O.V. Pikichian, eds), 120–134, Erevan: Izd. Akad. Nauk ArmSSR (in Russian).

Rogovtsov, N.N., 1990a: On the radiation field at a depth in a semi-infinite turbid medium containing a point directional source, *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana*, **26**, 1082–1088 (in Russian).

Rogovtsov, N.N., 1990b: On the asymptotics of the Green functions of the radiation transport equation for scattering media that have a cylindrical form, *Dokl. Akad. Nauk BSSR*, **34**, 1081–1084 (in Russian).

Rogovtsov, N.N., 1990c: On the solution of the Cauchy problem for infinite systems of differential equations that describe the process of a multiphoton collisionless excitation of molecules, *Differential Equations*, **26**, 436–441.

Rogovtsov, N.N., 1991a: Asymptotics of the Green function of the radiation transport equation for a plane layer that contains a unidirectional point source, *Dokl. Akad. Nauk BSSR*, **35**, 43–46 (in Russian).

Rogovtsov, N.N., 1991b: On integral characteristics of radiations fields in almost conservatively scattering homogeneous media, *Vesti Akad. Nauk BSSR, issue fiz. matem. navuk*, No. 4, 83–88 (in Russian).

Rogovtsov, N.N., 1991c: The radiative transfer in scattering media different configuration, in *Scattering and Absorption of Light in Natural and Turbid Media* (A.P. Ivanov, ed.), 58–81, Minsk: Institute Phiziki Akad. Nauk BSSR (in Russian).

Rogovtsov, N.N., 1992a: On asymptotical formulae of radiative transfer equations Green function for scattering media of a layer form or a rotation cylinder form, *Dokl. Akad. Nauk Belarus*, **36**, 26-29 (in Russian).

Rogovtsov, N.N., 1992b: On the derivation of the radiative transfer equation surface Green function for the case of cylindrical form medium, *Dokl. Akad. Nauk Belarus*, **36**, 598–601 (in Russian).

Rogovtsov, N.N., 1993: On integral relations for the non-stationary radiative transfer equation for arbitrary configuration scattering media, *Dokl. Akad. Nauk Belarus*, **37**, 34–38 (in Russian).

Rogovtsov, N.N., 1994: Radiative transfer in scattering and absorbing media of different configuration, Doctor of Sciences dissertation, Minsk: Institut Phiziki Acad. Nauk Belarus (in Russian).

Rogovtsov, N.N., 1996: On solving the characteristic equation of radiative transfer theory in an analytical form, in *Boundary-Value Problems, Special Functions and Fraction Calculus* (A.A. Kilbas, ed.), 305–312, Minsk, BGU (in Russian).

Rogovtsov, N.N., 1997: Analytical expressions for the radiation power coming from different configuration turbid media, *Dokl. Akad. Nauk Belarus*, **41**, 52–57 (in Russian).

Rogovtsov, N.N., 1999: *Invariance Properties and Principles, Application to Problems of Mathematical Physics*, Part 1, Minsk: Education Ministry of RB, BGPA (in Russian).

Rogovtsov, N.N., 2002: Inequalities and asymptotical formulae of radiation field characteristics in turbid media of non-concave form, *Dokl. NAN Belarus*, **46**, 50–53 (in Russian).

Rogovtsov, N.N., 2004: Application of invariance relations for solving boundary-value problems for ordinary differential equations, *Trudy Instituta Matematiki NAN Belarus, Minsk*, **12**, 145–149.

Rogovtsov, N.N., 2007: Application of invariance relations to deriving new integral equations for reflection and transmission functions of a plane-parallel turbid medium, *Teoreticheskaya i prikladnaya mekhanika, Minsk*, **22**, 72–77 (in Russian).

Rogovtsov, N.N., 2008: Relationship between solutions of families of two-point boundary-value problems and Cauchy problems, *Differential Equations*, **44**, 1–20.

Rogovtsov, N.N., Borovik, F.N., 1993: On solving the radiative transfer equations for a infinite homogeneous plane-parallel medium in a closed form, *Dokl. Akad. Nauk Belarusi*, **37**, 39–44 (in Russian).

Rogovtsov, N.N., Borovik, F.N., 2009: The characteristic equations of radiative transfer theory, in *Light Scattering Reviews*, vol. 4 (A.A. Kokhanovsky, ed.), 347–429, Chichester, UK: Springer–Praxis.

Rogovtsov, N.N., Karpuk, V.V., 2003: Analytical expression for mean and integral characteristics of the radiation fields in turbid media of different configuration, *Dokl. NAN Belarus*, **47**, 55–59 (in Russian).

Rogovtsov, N.N., Samson, A.M., 1975: The non-stationary radiation transfer in scattering media, *Astrophysics*, **11**, 293–302.

Rogovtsov, N.N., Samson, A.M., 1976: Search for the source function inside a plane-parallel scattering layer from the intensity of emerging radiation, *J. Applied Spectroscopy*, **25**, 1164–1169.

Rogovtsov, N.N., Samson, A.M., 1985a: Asymptotical expressions for the internal radiation fields in plane-parallel media and applications to calculation of mean emission durations of a layer and a spherical shell, *Astrophysics*, **23**, 468–476.

Rogovtsov, N.N., Samson, A.M., 1985b: Reduction of the problem on calculation of radiation fields in plane-parallel media coming arbitrary internal sources to the determination of radiance factors, Preprint No. 398, Institute Phiziki Akad. Nauk BSSR (in Russian).

Rogovtsov, N.N., Samson, A.M., 1987: On estimations of the mean number of photon scattering and the mean density of radiation in scattering media of non-concave form, *Dokl. Akad. Nauk BSSR*, **31**, 320–323 (in Russian).

Romanova, L.M., 1985: Radiation density and absorption in the homogeneous and horizontally inhomogeneous clouds, *Izv. Akad. Nauk SSSR, Fiz. Atmos. Okeana*, **21**, 830–840 (in Russian).

Rozenberg, G.V., 1977: Light ray – a cotribution to the theory of the light field, *Soviet Physics Uspekhi*, **20**, 55–80.

Sanchez, R., Ganapol, B.D., 1983: Benchmark value for monoenergetic neutron transport in one-dimension cylindrical geometry with linear anisotropic scattering, *Nucl. Sci. Eng.*, **84**, 61–67.

Sanchez, R., McCormick, N.J., 1982: A review of neutron transport approximation, *Nucl. Sci. Eng.*, **80**, 481–535.

Sanchez, R., McCormick, N.J., 2004: Discrete ordinates solution for highly forward peaked scattering, *Nucl. Sci. Eng.*, **147**, 249–274.

Santadrea, S. Sanchez, R., 2002: Positive linear and nonlinear surface characteristic scheme for the neutron transport equation in unstructured geometries, Proc. of Physor-2002, CD.

Shikhov, S.B., 1973: *Questions on Mathematical Theory of Reactors*, Moscow: Atomizdat (in Russian).

Shimizu, A., 1968: A calculation of the penetration of gamma rays through slabs by the method of invariant imbedding, *Nucl. Sci. and Eng.*, **32**, 184–194.

Shirkov, D.V., 1989: Invariance – automodelity – renormalization group, in *Invariance Principle and its Applications* (M.A. Mnatsakanian and O.V. Pickichian, eds), 457–473, Erevan: Izd. Akad. Nauk ArmSSR (in Russian).

Sobolev, V.V., 1951: A new method in light scattering theory, *Astronom. zhurnal*, **28**, 355–362 (in Russian).

Sobolev, V.V., 1956: Radiative transfer in a non-uniform medium, *Dokl. Akad. Nauk SSSR*, **111**, 1000–1003 (in Russian).

Sobolev, V.V., 1968: Diffusion of radiation in the medium large optical thickness for an anisotropic scattering, *Dokl. Akad. Nauk SSSR*, **179**, 41–44 (in Russian).

Sobolev, V.V., 1973a: Relationship between radiative and absorptive abilities of a medium for arbitrary scattering law, *Dokl. Akad. Nauk SSSR*, **212**, 1096–1098 (in Russian).

Sobolev, V.V., 1973b: Some application of the relationship between emissivity and absorptivity of medium, *Astrofizika*, **9**, 515–524 (in Russian).

Sobolev, V.V., 1975: *Light Scattering in Planetary Atmospheres*, New York: Pergamon Press.

Stokes, G.G., 1862: On the intensity of light reflected from or transmitted through a pile of plates, in *Mathematical and Physical Papers of Sir George Stokes*, 1904, vol. IV, 145–156, London: Cambridge University Press.

Sushkevich, T.A., et al., 1990: *Characteristic Method in Atmospheric Optics*, Moscow: Nauka (in Russian).

Tuchin, V.V., 1997: Light scattering study of tissues, *Uspehi fizicheskih nauk*, **167**, 517–539 (in Russian).

Tuchin, V.V., 2000: *Tissue Optics*, Bellingham: SPIE.

Ueno, S., 1957: The probabilistic method for problems of radiative transfer: II, Milne's problem, *Astrophys. J.*, **126**, 413–417.

Ueno, S., 1958: The probabilistic method for problems of radiative transfer: III, Line formation by coherent scattering, *J. Math. Mech.*, **7**, 629–641.

van de Hulst, H.C., 1963: A new look at multiple scattering, NASA Inst. for Space Studies. New York, Report.

van de Hulst, H.C., 1980: *Multiple Light Scattering. Tables, Formulas and Applications.* Vols. 1 and 2, New York: Academic Press.

Viik, T., 1982: A generalized principle of invariance and radiation field in multilayer atmospheres, *Astrophys. and Space Sci.*, **86**, 169–178.

Vladimirov, V.S., 1961: Mathematical problems of single-velocity theory of particle transfer, *Trudy MIAN SSSR*, **61**, (in Russian).

Vladimirov, V.S., 1971: *Mathematical Physics Equations*, Moscow: Nauka (in Russian).

Vladimirov, V.S., 1979: *Generalized Functions in Mathematical Physics*, Moscow: Nauka (in Russian).

Weyl, H., 1952: *Symmetry*, Princeton: Princeton University Press.

Weyl, H., 1968: *Gesammelte Abhandlungen*, Berlin: Springer.

Wigner, E.P., 1970: *Symmetries and Reflections*, Bloomington: Indiana University Press.

Yanovitskij, E.G., 1979: Invariance principle and integral relations for radiation fields in plane atmosphere, *Astron. zhurn.*, **56**, 833–844 (in Russian).

Yanovitskij, E.G., 1981: Anisotropic light scattering in an inhomogeneous atmosphere. Invariance principle and integral relations of the radiative transfer equations, *Astron. zhurn.*, **58**, 119–129 (in Russian).

Yanovitskij, E.G., 1991: Radiation flux in a weakly absorbing atmosphere of arbitrary optically thickness, *Izv. Akad. Nauk SSSR. Fiz. Atmos. Okeana*, **27**, 1241–1246 (in Russian).

Yanovitskij, E.G., 1997: *Light Scattering in an Inhomogeneous Atmosphere*, New York: Springer-Verlag.

Yengibarian, N.B., Melkonian, E.A., 1989: Radiative transfer a plane slab and Ambartsumian's equations, in *Invariance Principle and it Applications* (M.A. Mnatsakanian and O.V. Pikichian, eds), 326–333, Erevan: Izd. Akad. Nauk ArmSSR (in Russian).

Yengibarian, N.B., Mnatsakanian, M.A., 1974: On linear problems of transfer, *Dokl. Akad. Nauk SSSR*, **217**, 533–535 (in Russian).

Yi, H.C., Sanchez, R., McCormick, N.J., 1992: Bioluminescence estimation from ocean in situ irradiances, *Appl. Opt.*, **31**, 822–830.

Yosida, K., 1980: *Functional Analysis*, Berlin: Springer-Verlag.

Zege, E.P., Ivanov, A.P., and Katsev I.L., 1991: *Image Transfer through a Scattering Medium*, New York: Springer-Verlag.

Part III

Optical Properties of Bright Surfaces and Regoliths

8 Theoretical and observational techniques for estimating light scattering in first-year Arctic sea ice

Bonnie Light

8.1 Introduction

Earth's climate is known to be sensitive to the light scattering properties of sea ice across the Arctic Basin. For the purpose of understanding its interactions with shortwave radiation, sea ice can be described as a multiply scattering medium whose microstructural properties govern the quantity and quality of light scattering. The ice is populated with numerous inclusions of brine, air, precipitated salt crystals, and impurities embedded within a matrix of pure ice (e.g., see Weeks and Ackley, 1982). The multiple scattering caused by these inclusions in thick, bare sea ice generally causes more than half of the incident solar radiation to be backscattered to the atmosphere (Perovich, 1990). However, the structure of this composite material is complex and so the optical properties of sea ice have proven difficult to quantify and generalize. Since the liquid brine contained in sea ice must remain in freezing equilibrium with the ice, temperature changes can produce significant changes in the size and number distributions of these inclusions which determine the backscattering, absorption, and transmission of shortwave radiation at the frozen surface of the ocean.

The objective of this review is to summarize our quantitative understanding of the relationships between the structural and optical properties of sea ice. The work is based on direct laboratory observations of the response of both the ice microstructure and its optical properties to changes in temperature. Relationships between the ice microstructure and its inherent optical properties are defined and explored through the development of a structural-optical model for sea ice.

8.2 Background

The visual appearance of sea ice, from bright white to milky or translucent gray, is indicative of varying degrees of multiple scattering. Areas less than a square kilometer typically include ice types with varying age and thickness. Thick multiyear white ice has a relatively high albedo and small transmissivity, whereas thin, newly formed ice may backscatter much less light and transmit significant amounts of light to the ocean. Variations in ice thickness, age, and growing conditions promote

vertical gradients in salinity and temperature, which in turn cause vertical gradients in the size and number distribution of liquid and gas inclusions in the ice. Seasonal changes cause substantial temporal changes in the optical properties of sea ice as well. The onset of melt during the summer season produces dramatic changes in the optical properties when the snow melts and the highly reflective surface gives way to areas of bare melting ice, ponded ice, and a substantially thinner summer ice pack.

Direct measurements of inherent optical properties (IOPs), including the scattering coefficient (σ), scattering phase function (often represented by its asymmetry parameter, g), and absorption coefficient (κ), are difficult to make within multiply scattering media. One approach for obtaining these IOPs is to calculate them directly from information on the microstructure of the ice. The IOPs can then be incorporated into radiative transfer models and tested using measurements of apparent optical properties (AOPs).

A theoretical framework relating the physical properties of sea ice to the IOPs was laid out by Grenfell (1983, 1991). In this formulation, sea ice is modeled as a system consisting of pure ice, liquid brine inclusions, gas bubbles, precipitated salt crystals, and inclusions of biogenic or lithogenic origin. Absorption in sea ice is attributed to the pure ice matrix, the bulk brine volume fraction, and any particulate or biological inclusions present; scattering is assumed to be caused primarily by inclusions of brine, gas, and precipitated salts found within the ice. It is assumed that the individual inclusions scatter independently so that κ, σ, and g can be determined for each type of scatterer and combined incoherently.

Inclusions in sea ice occur over a wide range of scales. The largest brine inclusions may be centimeters long, whereas precipitated salt crystals can have edge lengths on the order of micrometers (Light, 1995). The size, refractive index contrast, and shape of an inclusion dictate how it scatters light. While these properties vary with inclusion type, they can also depend strongly on ice temperature. Changes in temperature produce changes in the size and salinity of brine inclusions. The refractive index of brine depends on its concentration (Maykut and Light, 1995), as does the precipitation or dissolution of a variety of solid salt crystals (Nelson and Thompson, 1954; Richardson, 1976; Marion et al., 1999).

Because natural sea ice typically exhibits large vertical and horizontal variations in structure, apparent optical properties measured in the field are usually not representative of ice with uniform IOPs, making it difficult to validate structural-optical models using field data. To address this problem, a laboratory study was developed to measure the optical and structural properties of sea ice under temperature-controlled conditions. In this review we present results from these studies, along with the development of a physically-based structural-optical model for first-year Arctic sea ice.

8.3 Approach

The foundation for this structural-optical model rests on laboratory experiments designed to quantify relationships between the structural and optical properties of natural sea ice. Fig. 8.1 shows a schematic of the approach. Structural and optical

data were collected concurrently in a temperature-controlled freezer laboratory using two isothermal samples of interior, first-year ice taken from adjacent locations in the same ice core. Varying the temperature of these samples produced large structural changes, allowing us to document the optical response of the ice under a wide range of conditions. Optical samples 5 cm in height were taken from the middle of the core where visual inspection indicated that the ice was relatively homogeneous. Observations of the microstructure (Fig. 8.1, A1) were made on vertical and horizontal thin sections using a high-resolution zoom lens and a black and white digital video camera. Numerous images were recorded, detailing the size and number density of brine tubes, brine pockets, gas bubbles, and to a limited extent, precipitated salt crystals. The various inclusions were counted and their size distributions parameterized. An equivalent spheres treatment (Grenfell and Warren, 1999) was applied and a temperature-dependent effective cross-sectional area ψ(T) estimated for each type of inclusion. Details of the microstructure observations are described in section 8.4 (see also Light et al., 2003a).

Concurrent AOP observations (Fig. 8.1, B1) were made on cylindrical core sections 10 cm in diameter and 5 cm thick. Transmitted, backscattered, and side scattered spectral radiances were recorded at temperatures between -1 and -35°C. Temperature was normally varied in 2–5°C increments, and the ice allowed to equilibrate for at least 24 hours before new measurements were made. Details of the optical observations are presented in section 8.5.

Optical measurements were interpreted using a two-dimensional Monte Carlo model (Fig. 8.1, B2) developed specifically to treat radiative transfer in multiply scattering domains with cylindrical geometry. The development and application of this model are summarized in section 8.6 (see Light et al., 2003b for full model description).

The structural-optical model (Fig. 8.1, A2) developed here is based on the theoretical framework of Grenfell (1983). The model uses direct observations of inclusion number densities and size distributions obtained from the microstructure observations to predict temperature and wavelength-dependent absorption coefficients $[\kappa(T, \lambda)]$, temperature-dependent scattering coefficients $[\sigma(T)]$, and temperature-dependent scattering asymmetry parameters $[g(T)]$ for brine inclusions, gas bub-

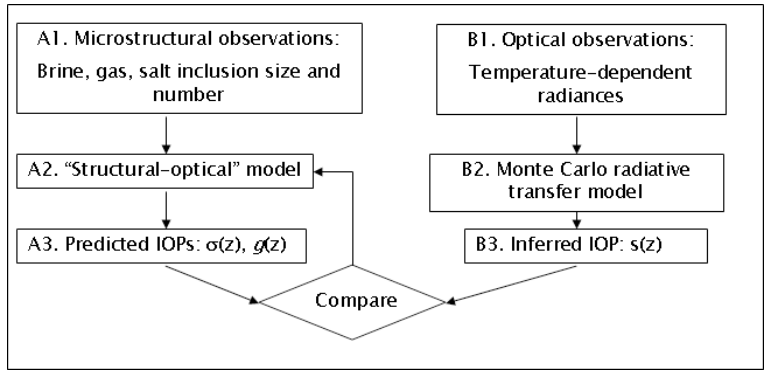


Fig. 8.1. Schematic of experimental approach.

bles, and precipitated salt crystals. Fundamental physical relationships between the ice structure and its optical properties are deduced from the AOP data (Fig. 8.1, B3) along with predictions from the structural-optical model (Fig. 8.1, A3). Section 8.7 details the development and refinement of the structural-optical model (see also Light et al., 2004).

8.4 Sea ice microstructure

This section presents a description of direct laboratory observations of sea ice microstructure that serve as the fundamental basis for estimating the scattering coefficients for the sea ice samples measured in this study.

8.4.1 Overview

Sea ice is distinguished from fresh water ice by the way it grows, its crystal characteristics, and its inclusions of liquid brine. These brine inclusions form when small ice bridges grow between advancing ice platelets (see Weeks and Ackley, 1982). Initially, the salinity of the brine trapped in the ice is close to that of the seawater from which the ice formed (typically 32–35 parts per thousand by mass; hereafter ‰). As the ice sheet grows thicker, the interior ice becomes more insulated from the warm ocean water and brine inclusions are subjected to progressively lower temperatures, causing the brine to become more concentrated to maintain its freezing-equilibrium salinity. To maintain equilibrium, water from the brine freezes to the surrounding ice lattice, reducing the size of the inclusions and causing the brine salinity to increase. As the brine concentrates, it becomes saturated with respect to certain salts which begin to precipitate. Conversely, increases in temperature cause ice to melt from the walls of brine inclusions, thus reducing brine salinity and causing the dissolution of solid salt crystals.

Cox and Weeks (1983) developed parameterizations for the bulk volumes of brine, gas, and precipitated salt as a function of ice temperature, density, and salinity. While knowledge of these equilibrium relationships is necessary for understanding how the physical properties of the ice respond to changes in temperature, detailed information about the size, number density, and spatial arrangement of these inclusions is also required for understanding how the optical properties of sea ice behave.

8.4.1.1 Brine inclusions

Measurements of brine inclusion size and number distributions have been made by imaging thin sections from both laboratory-grown and natural sea ice samples (e.g., Arcone et al., 1986; Eicken, 1993; Perovich and Gow, 1991; Perovich and Gow, 1996; Light, 1995; Cole and Shapiro, 1998; Eicken et al., 2000). Fig. 8.2 shows a photograph of a vertical thick section cut from a depth of approximately 80 cm in naturally grown first-year Arctic sea ice. In this low resolution view, both brine and gas inclusions are visible. A majority of the visible features are elongated brine inclusions, and while there appears to be considerable small-scale spatial

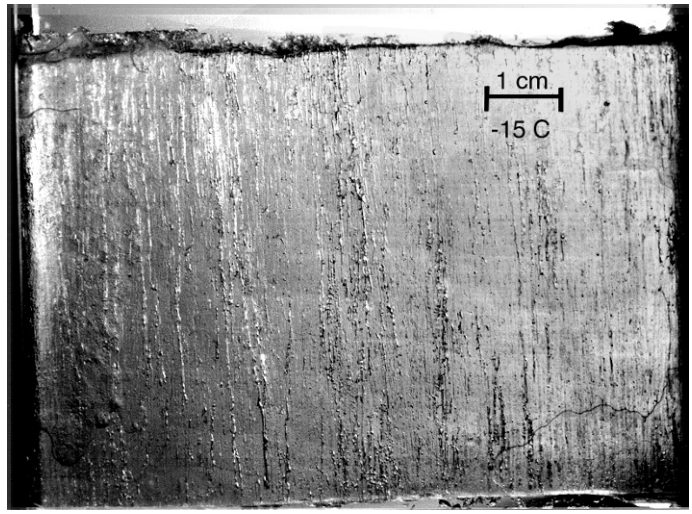


Fig. 8.2. Vertical thick section taken from interior first-year ice at a depth of approximately 80 cm. Sample thickness is approximately 5 mm.

variability, the overall structure is remarkably uniform. Observations from higher resolution imagery on horizontal thin sections from first-year sea ice (Perovich and Gow, 1996) indicate that average brine inclusion number densities range from 1.0 to 4.5 per mm^3 and cross-sectional areas of individual inclusions range from 0.0012 to 1.0 mm^2. Cole and Shapiro (1998), however, demonstrated that additional information about ice microstructure could be obtained by complementing the data from horizontal thin sections with high-resolution data from vertical thin sections. They found, for example, that brine inclusion shapes ranged from nearly spherical to elongated cylinders with vertical extent exceeding 15 times the diameter. While such studies have provided general information about brine inclusions, they do not address the role of gas bubbles and precipitated salts.

8.4.1.2 Gas bubbles

The density of sea ice increases above that of pure ice (0.917 g cm^{-3} at $-5°C$) as its salinity is increased, and decreases as its gas content is increased. The bulk density of first-year sea ice typically ranges between 0.89–0.93 g cm^{-3} (Perovich et al., 1998), where the lower end of this estimate corresponds to sea ice containing significant amounts of gas. Bubbles can be entrained into sea ice at the growth interface during the freezing process when gas dissolved in seawater comes out of solution. Void spaces can also form in ice above the freeboard level due to melt water drainage during the summer. In addition, gas bubbles are expected to form in brine inclusions as the ice warms and lower density ice melts into higher density liquid, forming a void within the inclusion. This suggests that there should be two primary types of bubbles in sea ice: those that occur within the ice lattice itself, and those that form within brine inclusions.

8.4.1.3 Precipitated salt crystals

As sea ice cools, brine in the liquid inclusions concentrates and becomes saturated with respect to certain salts. Observational and theoretical studies have been conducted to detail the species, precipitation temperatures, and total masses of salt that precipitate (Ringer, 1928; Gitterman, 1937; Nelson and Thompson, 1954; Richardson, 1976; Marion et al., 1999). The mass of precipitated salt is a unique function of temperature and ice salinity. The two most abundant salts are mirabilite ($Na_2SO_4 \cdot 10H_2O$) which begins to precipitate above $-8°C$, and hydrohalite ($NaCl \cdot 2H_2O$) which begins to precipitate at about $-23°C$. As shown in Fig. 8.3, the precipitation of each salt occurs gradually over a range of temperatures; the total mass of hydrohalite precipitated from seawater is more than 4 times that of mirabilite. These two hydrated salts are known to form monoclinic crystals, but there are no direct observations of their sizes or precipitation patterns within sea ice. Roedder (1984) suggested that hydrohalite crystals in fluid inclusions are commonly small ($\sim 1\,\mu$m). Although crystal sizes are expected to range over some distribution, at this point we only have estimates of effective crystal sizes in sea ice that are based on comparison of observed and modeled optical properties. Light (1995) utilized this method to estimate an effective edge length of 8–9 μm for precipitated salt crystals in laboratory-grown sea ice.

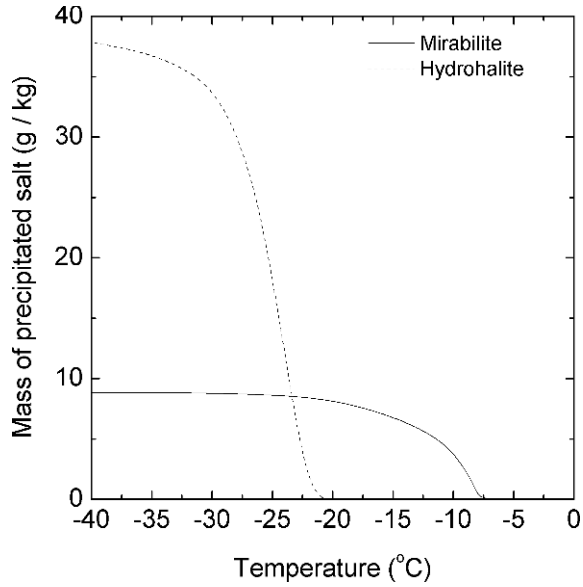


Fig. 8.3. Mass of mirabilite and hydrohalite precipitated from a kilogram of freezing seawater as a function of temperature. Salinity is 35‰. Data are from Richardson (1976).

8.4.2 Laboratory observations

While the previous observations provide a general picture of the microstructure of sea ice, many aspects of the structure which could affect light scattering remain uncertain due to the limited spatial resolution used in these earlier studies. The highest resolution study (Perovich and Gow, 1996), for example, could only detect inclusions larger than 0.05 mm. Many of the earlier data were also obtained from horizontal thin sections (e.g., Perovich and Gow, 1991, 1996; Eicken, 1993: Light, 1995), restricting the information to inclusion cross-sections and masking the three-dimensional nature of the microstructure. In addition, there is little information about how inclusions interact as temperature changes. Of particular interest are the formation and growth of bubbles within brine inclusions, the merging and coalescing of brine inclusions at higher temperatures, and the precipitation characteristics of salt crystals. Such information is needed in the formulation of a more comprehensive structural-optical model for sea ice.

To address these shortcomings, a laboratory experiment was designed to obtain very high resolution imagery of natural sea ice samples under a wide range of temperatures, the focus being on first-year ice. Because of the need for accurate temperature control, measurements were carried out in a laboratory cold room, rather than in situ. Ice cores used in this study were extracted from shorefast ice near Point Barrow, Alaska in May 1994. The ice was 1.65– 1.75 m thick. The temperature was about −5.5°C at the ice surface and increased linearly with depth to the freezing point at the base of the ice. Measured ice salinity profiles were typically C-shaped with values ranging from 7–9‰ near the surface, 4–5‰ in the interior, and about 10‰ near the bottom. Calculated brine volumes ranged from 5–12 %. Gas volumes were about 4 % in the upper 10 cm of the ice, but less than 1% below this (see Perovich et al., 1998). Cores, 10 cm in diameter, were removed from the upper meter of the ice, shipped to Seattle in dry ice, and stored in a laboratory freezer at the University of Washington at −20°C. While it is not known how shipping core samples in dry ice and subsequent long-term storage affect the ice microstructure, we proceeded to analyze these samples for their microstructural and optical properties.

Cores were cut into thin sections for laboratory examination. Samples were taken from the lower portions of the cores, corresponding to a depth of 60–80 cm in the ice. These samples had relatively small gas volumes and brine inclusions which tended to be isolated, unlike ice near the growth interface or upper surface. Because the vertical dimension of brine inclusions is often much larger than their horizontal dimension, vertically oriented, 8×8 cm slabs were cut from the cores and microtomed to a thickness of approximately 2 mm. The sections were prepared at −15°C to minimize brine drainage. Thin sections were sealed between clear glass plates which allowed them to be mounted vertically for monitoring the thermal evolution of the ice. The sealed glass plates kept the structure intact, even at high temperatures where the porosity was high and the samples extremely fragile. Samples were mounted on an x-y translation stage, allowing them to be moved as much as 3 cm laterally and 3 cm vertically, and then returned to a predetermined position. This made it simple to revisit and photograph specific scenes in the ice as conditions changed. Image sizes were calibrated using a grid with ruled lines separated by 50 micrometers.

To record the microstructure, a black and white CCD video camera (Panasonic model WV-BP500) with 682 by 492 pixel resolution and a high magnification zoom lens (Leica MonoZoom 7 Optical System) was used. The high magnification lens (coupled with a 3X amplifier) gave us the ability to resolve inclusions as small as 0.01 mm, considerably smaller than in previous studies. A majority of our images were made with the zoom lens set to record approximately 0.003 mm per pixel, while some higher resolution images were made at 0.002 mm per pixel. Our objective, however, was to monitor features and processes that are important to understanding the optical properties of the ice, not necessarily to document the smallest features possible. Images obtained by the CCD camera were fed to a high resolution monitor visible through the cold room window. This monitor was used to select and focus images. Images were also fed to a MacIntosh Power PC, equipped with a video frame grabber (Scion VG-5) and image processing software (NIH Image 1.57). To facilitate alignment and reduce the effects of vibrations, components in the cold room were mounted on a 2 m long optical rail (see Fig. 8.4).

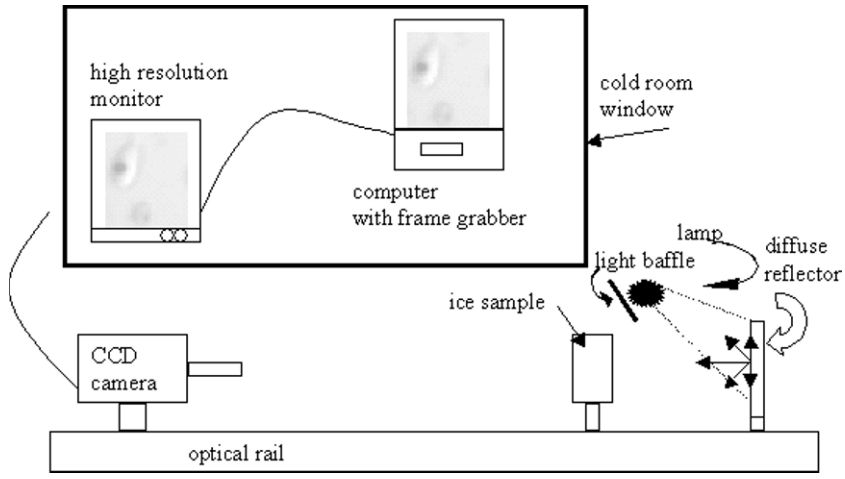


Fig. 8.4. Schematic of imaging system used for observing ice microstructure. The optical rail, sample, light source, diffusely reflecting target, and CCD camera are in the cold room. The focusing monitor and computer with frame grabber are outside of the cold room.

Information on crystallography is best obtained from images recorded in polarized light, whereas inclusions of brine and gas are best viewed in transmitted or reflected natural light. Experiments with both types of illumination indicated that high contrast images with uniform lighting could be obtained with transmitted light from a diffuse source. To achieve this, an incandescent light was aimed at a diffusely reflecting Spectralon® panel mounted directly behind the sample. The source was baffled to prevent direct radiation from illuminating or heating the sample.

8.4.3 Microstructure at −15°C

Structural observations to determine size distributions for brine inclusions and gas bubbles in cold ice were initially made at −15°C where changes in microstructure are not highly sensitive to changes in temperature. About 100 images were recorded from three different thin sections. Fig. 8.5 shows photomicrographs of one sample illuminated in transmitted light. On the left is a mosaic composed of 30 individual images. After registering the 30 images, the dimensions of the composite scene are approximately 4.7 × 12.1 mm. Several types of features stand out distinctly:

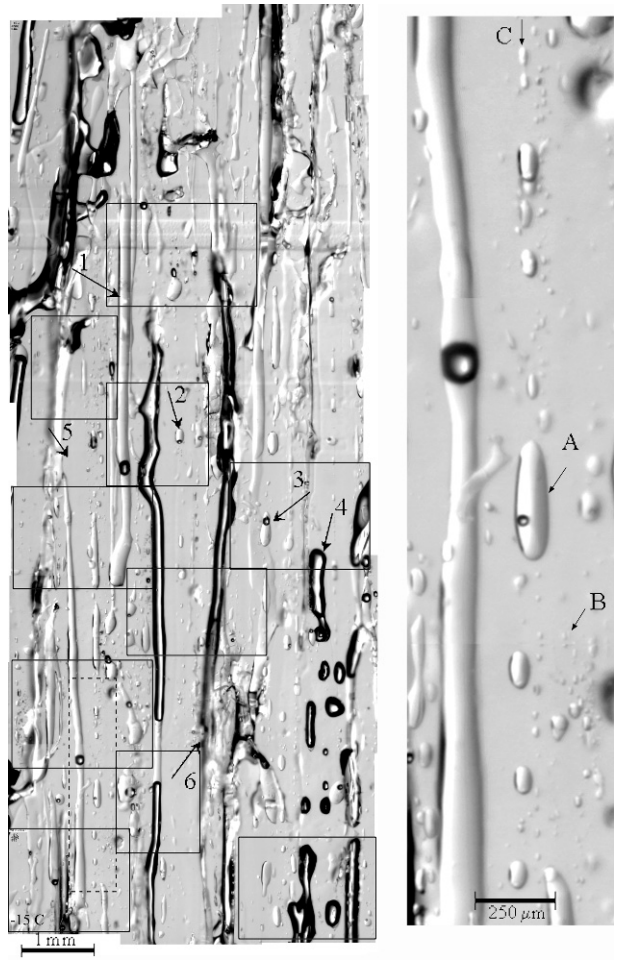


Fig. 8.5. Photomosaic of vertical thin section of first-year ice in transmitted light at −15°C. Ten boxed subregions were used for counting inclusions. Overall dimensions of scene are 12.1 × 4.7 mm, with 2 mm thickness. Arrows indicate examples of (1) brine tubes, (2) brine pockets, (3) bubbles, (4) drained inclusions, (5) transparent areas, and (6) poorly defined inclusions. (right) An enlargement of box outlined with dashed line. Arrows indicate (A) solitary brine pocket, (B) cluster of small pockets, and (C) string of pockets.

elongated brine tubes (arrow 1), smaller, isolated brine pockets (arrow 2), and gas bubbles within brine inclusions (arrow 3). Brine features that appear to have excessively high contrast were considered to be inclusions that were cut open and drained during sample preparation (e.g., arrow 4). Gas bubbles within brine pockets have similarly high contrast but can usually be discerned by their nearly spherical shape. There are also areas where no inclusions exist in the ice, and the ice appears highly transparent (arrow 5). The horizontal stripes across the image are scratches produced by the microtome blade. Microtome cuts were intentionally made across the sample (rather than down it) to avoid creating artifacts that might interfere with resolving features in the vertically oriented structure.

8.4.3.1 Brine inclusions

Data were collected on the maximum horizontal and vertical dimensions of brine inclusions visible within the ten boxed areas shown in Fig. 8.5. These areas were selected for sampling because they appeared to be representative of the overall structure. The sample volume within these boxes was 48.8 mm^3 and more than 1600 inclusions were counted. The sizes of individual inclusions ranged from 0.01 mm to 8.0 mm in length (l), and 0.01 mm to 0.230 mm in diameter. Inclusions with l smaller than 0.50 mm were arbitrarily classified as brine pockets. Examples of these features are shown in the enlargement on the right-hand side of Fig. 8.5. The large brine pocket (arrow A) has length 0.40 mm, diameter 0.11 mm, and a length-to-diameter aspect ratio (γ) of 3.6. Numerous small pockets are also visible in this same image; for example, arrow B points to an area of concentrated pockets whose recorded length and diameter are 0.01 mm with $\gamma = 1.0$. Pockets frequently appeared in clusters (arrow B) or in vertical strings (arrow C), and were always oriented with their long dimension in the vertical. Inclusions with $l \geq 0.50$ mm were classified as brine tubes. All observed tubes had near-vertical orientation. Since tubes were uncommon compared to pockets, the entire mosaic (sample volume 84.5 mm^3) was used to estimate their number density.

The observed size distribution of the brine inclusions is shown in Fig. 8.6. The data were binned by inclusion length and the number density, $N(l)$, was adjusted to account for the width of each bin. The data are well represented ($r^2 = 0.92$) by a power law:

$$N(l) = 0.28 l^{-1.96}, \tag{8.1}$$

where 0.01 mm $\leq l \leq$ 8 mm. Integrating Eq. (8.1) between these limits yields a total number density of 24 brine inclusions per mm^3. This appears to be representative for the sample as a whole, although individual boxes in Fig. 8.5 had number densities as large as 50 inclusions per mm^3. There is also a clear dependence of the aspect ratio on l as shown in Fig. 8.7. This relationship can also be approximated ($r^2 = 0.77$) by a power law:

$$\gamma(l) = 10.3 l^{0.67}, \tag{8.2}$$

where $1 \leq \gamma \leq 70$ and $l > 0.03$ mm. Nearly all the 1244 brine pockets observed to be less than 0.03 mm in diameter were recorded as spherical ($\gamma = 1$) with $l = 0.01$ mm, and are indicated by a single point in Fig. 8.7. Because inclusions this size were near the limit of our resolution, we assume that $\gamma = 1$ for all inclusions with $l < 0.03$ mm.

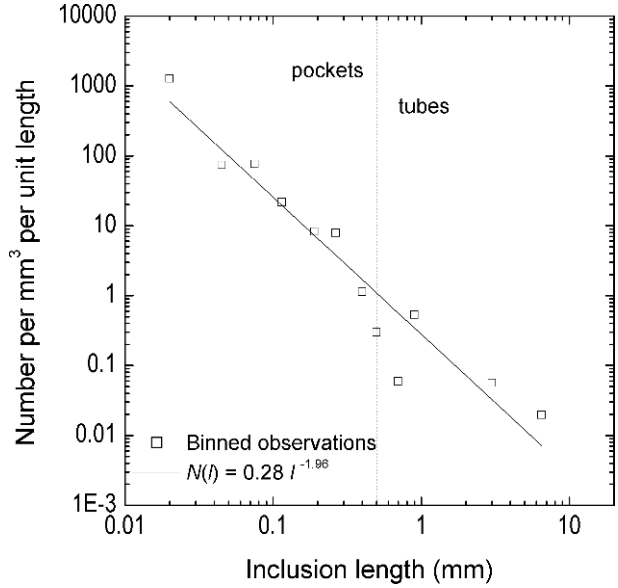


Fig. 8.6. Observed size distribution for brine inclusions at -15°C as a function of inclusion length. The solid line corresponds to a power law fit ($r^2 = 0.92$), and the dotted line divides pockets from tubes.

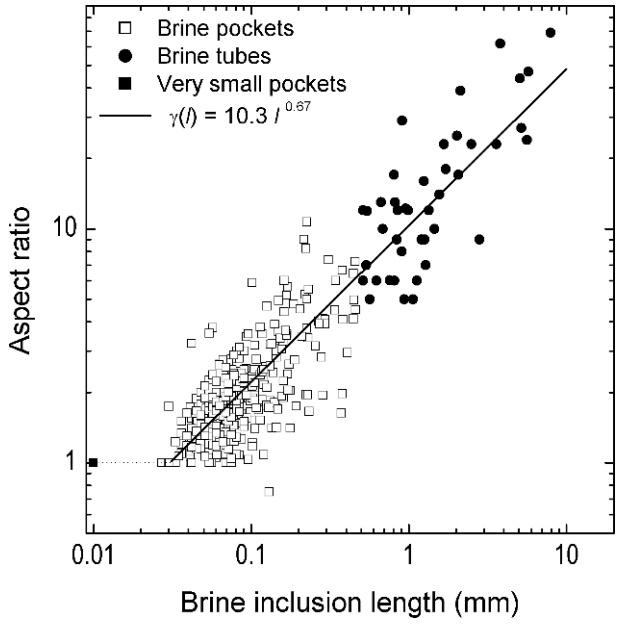


Fig. 8.7. Vertical-to-horizontal aspect ratio as a function of inclusion length for observed brine inclusions at -15°C. The solid line corresponds to a power law fit ($r^2 = 0.77$), and the dotted line indicates the parameterization for small inclusions used in this study. The solid square represents the numerous small inclusions that were at the resolution limit; open squares represent brine pockets, and solid circles represent brine tubes.

These results can be compared with the observations of Perovich and Gow (1996) and Cole and Shapiro (1998), both of whom studied ice extracted at the same time and location as the cores used here. Cole and Shapiro measured brine inclusion sizes in both vertical and horizontal thin sections taken from interior ice at a depth of 0.8 m. They observed inclusions with an average diameter of 0.27 mm in the horizontal plane and an average length of 2.4 mm in the vertical, values within the range we observed. However, they were not able to resolve individual inclusions smaller than 0.1 mm and did not estimate the size distribution or average number density. Perovich and Gow (1996) reported size distributions based on cumulative distributions of inclusion cross sections derived from horizontal thin sections. To compare our results with those of Perovich and Gow, we estimated horizontal cross-sectional areas of the inclusions in our images using the maximum observed diameter for each inclusion. Because Perovich and Gow observed mean horizontal axis ratios of 1:4 in their observations of inclusions with cross-sectional area greater than 0.002 mm^2, we examined imagery of a horizontal cross-section cut adjacent to the vertical thin section shown in Fig. 8.5. While this imagery was not examined for inclusion size and number distribution, it was scrutinized for information about the horizontal aspect ratio of the inclusions. While many of the inclusions did not have circular cross-section, there was no clear indication that horizontal aspect ratios were elongated in a particular direction. As a result, brine inclusions were initially assumed to have average horizontal aspect ratios equal to unity. Our average number density of 24 per mm^3 is 15 times larger than the value of 1.6 per mm^3 observed by Perovich and Gow. This difference is due primarily to the detection of smaller inclusions in our images.

Besides the absence of small inclusions in the Perovich and Gow observations, there are other differences between the two size distributions. Perovich and Gow found cross-sectional areas up to 1 mm^2 which correspond to inclusion diameters of 1.1 mm. In our study, the largest diameter measured was 0.23 mm (corresponding to an area of 0.04 mm^2). Number densities in our sample were as much as a factor of 2 smaller where the two distributions overlap (0.01–0.04 mm^2). We suspect our sampling technique may have selectively avoided larger tubes and know that it avoided brine channels because vertical thin sections are difficult to keep intact when such features are present. Perovich and Gow sampled a volume of 282 mm^3, more than three times greater than our sample volume. It is, therefore, reasonable to expect that their results may better represent the number density of larger brine tubes in first-year ice. But it should also be noted that the Perovich and Gow sample was imaged at $-5.7°C$ and had a brine volume of 5.5%, while our samples were measured at $-15°C$ and had a visible brine volume of only 1.2%. This visible brine volume was calculated by assuming that all the inclusions were isometric in the horizontal plane. Because the Perovich and Gow sample had such large brine volume, we might expect that their number density curve would shift to somewhat smaller cross-sectional areas if their measurements had been made at $-15°C$, bringing it into closer agreement with our observations.

Fig. 8.8 shows the relative volume of visible brine in our sample for several inclusion size categories. The bin widths are arbitrary, however, the smallest bin represents brine pockets that were below the resolution of the Perovich and Gow study; the smallest three bins include all the pockets, while the largest three bins

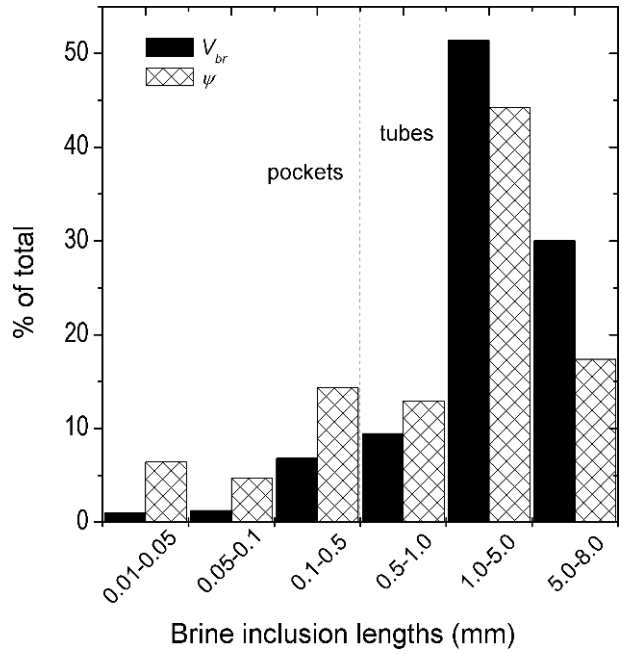


Fig. 8.8. Percentages of brine volume (V_{brine}) and equivalent cross-sectional area (ψ) for several size categories of brine inclusions at $-15°$C. Bin widths are arbitrary, but the distinction between pockets and tubes is indicated. The smallest bin represents pockets not resolved in the Perovich and Gow (1996) study.

represent the tubes. While they are the most numerous (79% by number), inclusions in the smallest bin contain only a small fraction of the total brine volume. Although tubes contained about 90% of the brine seen in the thin section, there is reason to believe that they may account for an even larger percentage of brine volume in the ice as a whole. The ice from which the thin section was extracted had an average salinity of 4.7‰ and a density of $0.915\,\mathrm{g\,cm^{-3}}$, resulting in a predicted brine volume of 1.9% at $-15°$C, significantly larger than the total visible brine volume of 1.2%.

It is likely that the missing brine was contained either in larger brine tubes that existed in the core sample but not in the thin section, or within visible inclusions whose volume was underestimated as a result of the assumption of horizontal isotropy. The supposition that larger tubes existed in the core sample is consistent with Perovich and Gow's size distribution which shows inclusions with large cross-sectional areas that were not present in our thin sections. Extrapolating our size distribution to include tubes up to 15 mm in length would produce a total brine volume of 1.9%. In this case, 94% of the total brine volume would be in tubes and only 6% in pockets. If, instead of extending the distribution out to include longer tubes, existing tubes are taken to be anisometric in the horizontal plane, the missing brine can also be accounted for. If a horizontal aspect ratio of 1.64 and a maximum inclusion length of 8 mm are assumed, the observed brine volume would also be 1.9%. In all likelihood, both factors are probably involved, but it

is difficult to assess their relative importance without further detailed study of additional sample material.

8.4.3.2 Equivalent cross-sectional area

Basic objectives of the structural observations were: (i) to obtain data which could be used to develop more accurate predictions of light absorption and scattering in sea ice, and (ii) to determine the optical importance of different elements in the microstructure. To make a quantitative comparison of the effects of various types of inclusions on radiative transfer, we represent all inclusions as equivalent spheres and calculate their total equivalent cross-sectional area (ψ):

$$\psi = \pi \int_{l_1}^{l_2} r_{eq}^2(l) N_{eq}(l)\, dl, \tag{8.3}$$

where ψ is per unit length, $r_{eq}(l)$ is the equivalent spherical radius for inclusions of length l, and $N_{eq}(l)$ is the corresponding number of equivalent spheres per unit volume. The optical scattering coefficient is proportional to ψ when inclusion sizes are much larger than the wavelength of the light. The scattering coefficient specifies how much light is scattered within a domain, but it does not specify the degree to which scattering is forward or backward directed. Redirection of light is tied to the scattering phase function and depends on the size, shape, and relative refraction of scatterers.

A variety of approaches has been applied to the calculation of equivalent spheres. When the total volume of all inclusions is conserved, absorption is accurately represented. However, scattering is generally underestimated because spheres of equal volume have less surface area than the original ellipsoids or cylinders. To accurately represent the scattering, the total surface area must also be conserved. According to Grenfell and Warren (1999), a collection of randomly oriented cylindrical inclusions can be represented by a collection of equivalent spheres with radius r_{eq} and number density N_{eq} whose total surface area and volume are the same as the original population. Fig. 8.9(a) shows calculated r_{eq} values as a function of observed inclusion length and aspect ratio for the three different conservation schemes. All r_{eq} are about the same for small brine pockets since they are approximately spherical to begin with. As inclusion length increases, conservation of both total surface area and volume produces substantially smaller r_{eq} but larger N_{eq} values than are obtained by conserving surface area alone. The value of r_{eq} decreases as it becomes necessary to increase the number of equivalent spheres per inclusion to conserve both. The discontinuity in $r_{eq}(l)$ arises because prolate ellipsoids with major axis l were used to calculate the volume and surface area for pockets ($l < 0.5\,\mathrm{mm}$), and cylinders with height l were used to represent the tubes. In Fig. 8.9(b), values of N_{eq} per inclusion show that when r_{eq} is approximately equal to $l/2$, which holds for $l \leq 0.03\,\mathrm{mm}$, there is approximately one equivalent sphere per inclusion. As inclusions increase in size and aspect ratio, the number of equivalent spheres per inclusion increases to its maximum value of 18 for $l = 8\,\mathrm{mm}$. Values of ψ reported below will always be calculated conserving both volume and surface area.

When integrated over all brine pockets ($l = 0.01$–$0.5\,\mathrm{mm}$), Eq. (8.3) predicts that $\psi = 30\,\mathrm{m}^{-1}$; similarly, $\psi = 80\,\mathrm{m}^{-1}$ for brine tubes ($l = 0.5$–$8\,\mathrm{mm}$), giving a

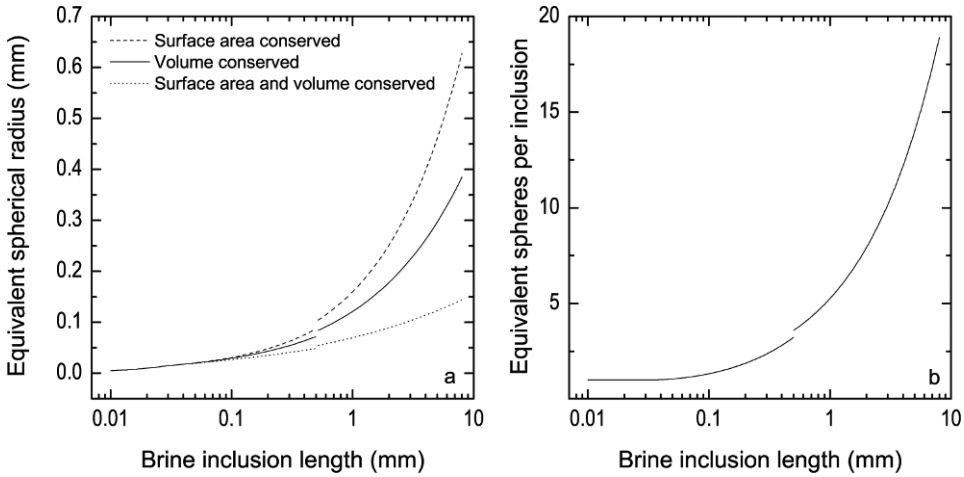


Fig. 8.9. (a) Equivalent spherical radius as a function of inclusion length as computed using equal surface area, equal volume, and simultaneous equal surface area and volume methods for the empirical fits describing lengths and aspect ratios observed in this study. (b) Number of equivalent spheres per inclusion for inclusion lengths and aspect ratios observed in this study. The discontinuities at $l = 0.50\,\mathrm{mm}$ result from treating brine pockets as prolate ellipsoids and tubes as cylinders.

total $\psi = 110\,\mathrm{m}^{-1}$ for all brine inclusions. Contributions made to ψ by various subcategories are shown in Fig. 8.8. While ψ due to the tubes is larger, the contribution from the pockets is significant. For example, even though only 10% of the brine volume is contained in pockets, it appears from this analysis that these pockets account for 25% of the light scattering due to the brine inclusions observed at $-15°\mathrm{C}$. These estimates are sensitive to the assumption that brine inclusions are isometric in the horizontal plane. If the horizontal asymmetry parameter were increased to the value (1.64) needed to explain the brine volume of the bulk core sample, the estimated value of ψ for brine tubes alone would increase 37% to $110\,\mathrm{m}^{-1}$.

8.4.3.3 Gas bubbles

The boxed areas in Fig. 8.5 were also used to estimate the size distribution of gas bubbles. All of the observed bubbles (about 60) were nearly spherical and all were contained within brine inclusions; none were observed in the ice itself. Gas bubble radius (r_{gb}) ranged from 0.004 to 0.07 mm at $-15°\mathrm{C}$. Fig. 8.10 shows the observed distribution as a function of r_{gb}; also shown are results from three other studies. As with the brine inclusions, a power law provides a good fit ($r^2 = 0.94$) to the observed distribution in this study:

$$N(r_{gb}) = 0.06(r_{gb})^{-1.5}. \tag{8.4}$$

When integrated over the observed size limits, the total number density was 1.3 per mm^3. This number density is only 5% of the number density for brine inclusions, indicating that while all bubbles were observed to be within brine inclusions, such

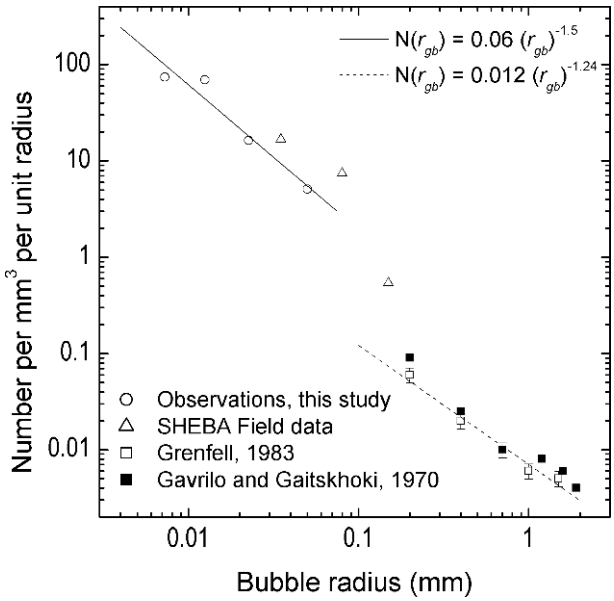


Fig. 8.10. Observed gas bubble size distributions for first-year ice in this study, first-year ice at the SHEBA field site, freezing lead ice (Grenfell, 1983), and 'bubbly' ice (Gavrilo and Gaitskhoki, 1970). Power law fits are also shown for the data in this study ($r^2 = 0.94$, solid line) and for Grenfell (1983) (dashed line). Here $r_{\rm gb}$ is the gas bubble radius, and $N(r_{\rm gb})$ is the gas bubble number density.

compound inclusions were relatively rare compared to brine inclusions with no bubble.

Since the bubbles are mostly spherical, use of equivalent spheres is greatly simplified and the effective radius is generally equal to the observed radius. The value of ψ_{gb} integrated between the observed bubble size limits was $2\,\mathrm{m}^{-1}$, only about 3% of the integrated value for all brine inclusions.

8.4.3.4 Mirabilite crystals

Mirabilite crystals were frequently seen at temperatures below $-8°$C (e.g., Fig. 8.11). Crystals were observed in piles at the bottom of brine tubes (arrow 1), in clusters strewn throughout the tubes (arrow 2), and stuck at narrow pinches in tubes (arrow 3). When piled at the bottom of brine tubes, the crystals often appeared to have tunneled downwards, extending the bottom of the tube. Mirabilite crystals were rarely observed in smaller brine pockets, presumably because the crystals were too small to resolve. The largest crystal in Fig. 8.11(b) has a diameter of 0.14 mm while the smallest measured diameter was 0.015 mm. The crystals tended to have rounded edges and irregular shapes, and may have coarsened after initial precipitation. Light (1995) inferred an effective crystal size of 0.009 mm for mirabilite, and this supports the conclusion that crystals in most brine pockets were too small to resolve. The smallest brine inclusion that could grow a single 0.01 mm mirabilite crystal would necessarily have a volume corresponding to $l \geq 0.06$ mm,

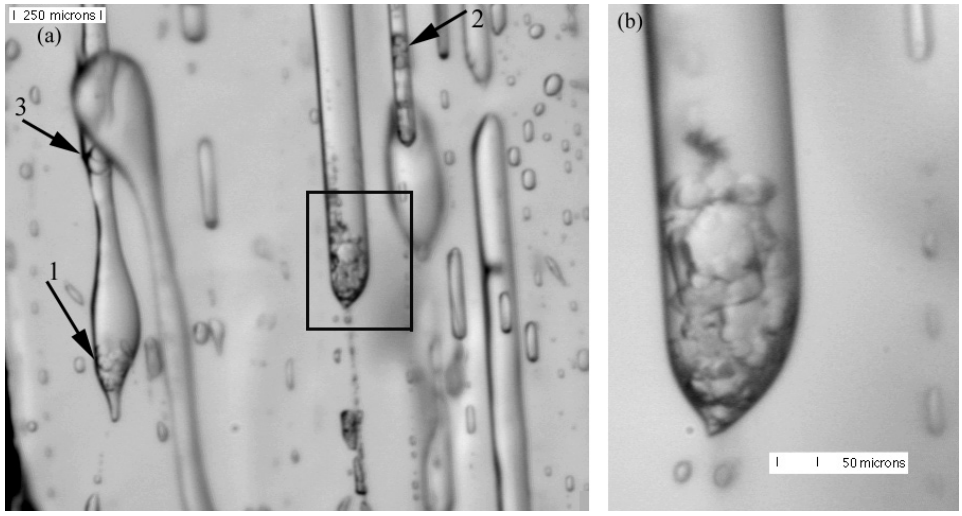


Fig. 8.11. (a) Photomicrograph of thin section showing individual mirabilite crystals at −15°C. (b) Enlargement of box in (a).

based on our observations of $\gamma(l)$. Because there were many brine inclusions with $l < 0.06$ mm, we believe there must be many crystals with edge length <0.01 mm. In addition, the refractive index contrast of mirabilite crystals relative to brine is smaller than that for brine inclusions in ice, so a brine inclusion with $l = 0.01$ mm may be visible in an image while a mirabilite crystal with the same size may not be. Temperature-dependent relative refractive indexes of the various types of inclusions responsible for light scattering in sea ice are presented in Table 8.1.

To estimate an effective optical cross-sectional area for mirabilite, we followed Light (1995) and simply assumed an effective crystal edge length of 0.01 mm since a direct estimate of the size distribution was not feasible. We also assumed a sample salinity of 3‰ commensurate with the visible brine volume (1.2%) of the thin section at −15°C. For this salinity, the predicted mass fraction of precipitate at −15°C is $0.58\,\mathrm{g\,kg^{-1}}$. Taking a density of $1.464\,\mathrm{g\,cm^{-3}}$ for mirabilite (Porter and Spiller, 1956), along with the monoclinic characteristics of the crystals, we estimated the total number density of crystals to be approximately $270\,\mathrm{mm^{-3}}$ at −15°C. Application of the equivalent spherical treatment to the mirabilite crystals at −15°C yielded $\psi = 50\,\mathrm{m^{-1}}$. This value is larger than that for brine pockets, about 60% of the value for brine tubes, and much larger than ψ for gas bubbles.

8.4.4 Temperature-dependent changes

While microstructural data described in the previous sections should allow the calculation of equivalent cross-sectional areas and, ultimately, scattering coefficients for first-year sea ice at −15°C, a general treatment should also be able to describe how these properties evolve with temperature. When the ice cools, the laws of freezing equilibrium predict that brine pockets and gas bubbles will shrink, mirabilite crystals will begin to precipitate below −8°C and hydrohalite crystals precipitate

Table 8.1. Temperature-dependent refractive index (n) for brine and relative refractive indexes (m) for the various inclusion types in sea ice. All values are for wavelength 589 nm.

Temperature (°C)	n for brine	m brine relative to ice	m air relative to ice	m air relative to brine	m mirabilite relative to brine	m hydrohalite relative to brine
−32	1.400	1.068	0.763	0.714	0.997	1.022
−30	1.397	1.067	0.763	0.716	0.999	1.023
−28	1.394	1.064	0.763	0.717	1.001	1.026
−26	1.390	1.061	0.763	0.719	1.004	1.029
−24	1.386	1.058	0.763	0.721	1.007	1.032
−22	1.383	1.055	0.763	0.723	1.010	1.034
−20	1.380	1.053	0.763	0.725	1.012	1.036
−18	1.376	1.051	0.763	0.727	1.014	1.039
−16	1.373	1.048	0.763	0.728	1.017	1.041
−14	1.370	1.046	0.763	0.730	1.019	1.044
−12	1.367	1.043	0.763	0.732	1.022	1.046
−10	1.363	1.040	0.763	0.734	1.024	1.049
−8	1.358	1.037	0.763	0.736	1.028	1.053
−6	1.354	1.033	0.763	0.739	1.031	1.056
−4	1.348	1.029	0.763	0.742	1.035	1.061
−2	1.342	1.024	0.763	0.745	1.041	1.066

at −23°C. When warmed, brine pockets and gas bubbles enlarge, and precipitated salts gradually dissolve. However, these laws do not predict how the size distributions of the various inclusions change or how these inclusions interact, factors which are likely to affect how the optical properties of the ice respond to changes in temperature. To obtain such information, ice samples were first cooled from −15 to −30°C, then warmed to −2°C. Changes in microstructure were monitored at 2–5°C intervals. The samples were cooled first, then warmed, to prevent the irreversible changes that occur when the ice becomes very warm from affecting the observations. Limitations in temperature regulation of the cold room produced a ±1°C uncertainty in actual sample temperatures. While the size of brine and gas inclusions in thin sections was observed to adjust rapidly (on the order of minutes) to changes in temperature, it is not clear whether salt precipitation and dissolution occurred as quickly. Rapid warming experiments (Adams and Gibson, 1930) suggest that these processes may require a few hours to reach equilibrium. For this reason, samples were held at each temperature for a minimum of 24 hours before measurements were taken.

8.4.4.1 Cooling sequence

Fig. 8.12 shows an example of structural changes observed in a 2.0 × 1.5 mm area of ice as it was cooled from (a) −15°C to (b) −20°C to (c) −25°C. The area contained three large brine tubes and numerous brine pockets. At −15°C, two of the tubes were observed to contain bubbles (arrows 1 and 2), as were several of the smaller pockets. The tube in the center of the image also contained several large crystals of mirabilite that had settled on top of the bubble (arrow 3). As the

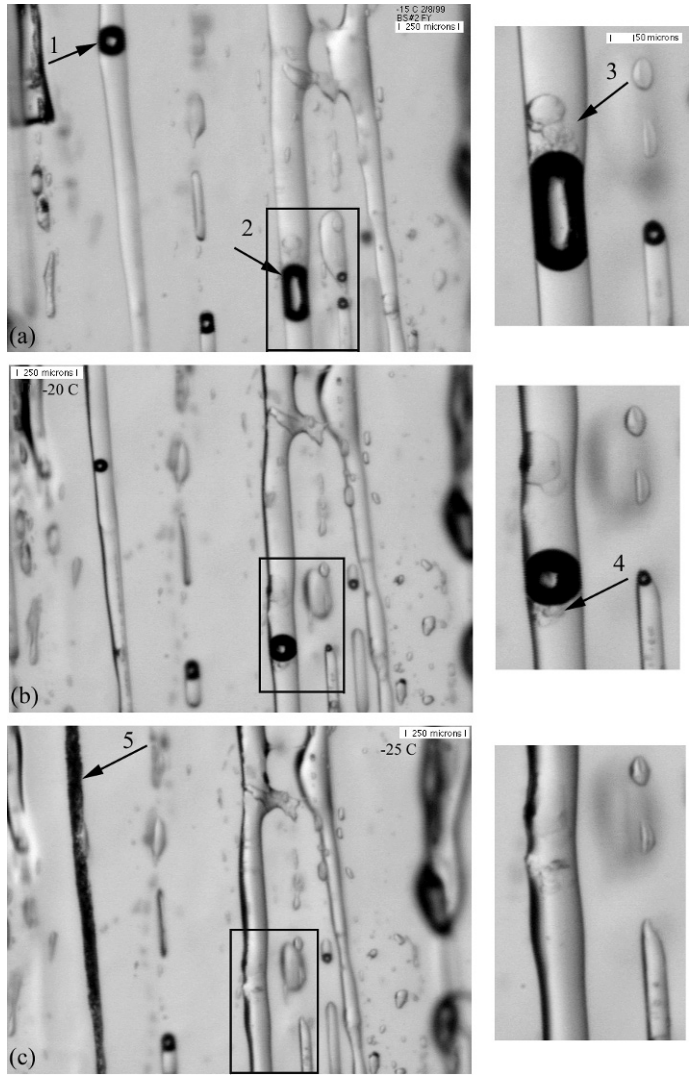


Fig. 8.12. Sequence showing changes in microstructure as the ice is cooled from (a) −15°C to (b) −20°C to (c) −25°C; (right) high magnification views of the (left) boxed areas in the larger images. Brine tubes, pockets, bubbles, and mirabilite crystals are observed to undergo significant evolution as the temperature changes. Arrows 1 and 2 point to bubbles within tubes, arrows 3 and 4 point to mirabilite crystals, and arrow 5 indicates a tube that has likely filled with hydrohalite slush.

sample was cooled to −20°C, each of the brine and gas inclusions shrunk in size. In fact, some bubbles disappeared altogether. All the tubes decreased in diameter. Pockets generally shrank in size while maintaining approximately constant γ. At −20°C, freezing equilibrium relationships predict that each brine inclusion should have about 82% of its volume at −15°C; at −25°C, this volume drops to 42%. If the tubes are assumed to be azimuthally symmetric and to have constant length,

the observed decreases are to 81% at −20°C and 57% at −25°C; 57% is the volume predicted at −24°C and it is likely that the temperature of the ice was closer to this value at the time the bottom photograph was taken. Note that the ice-brine refractive boundaries appear more distinct because the refractive index (n) of the brine increases from 1.370 at −15°C to 1.379 at −20°C (Maykut and Light, 1995). This change increases the contrast between brine and ice (n_{ice} = 1.30 to 1.32 at visible wavelengths), but decreases the contrast between mirabilite crystals (n = 1.396) and brine.

The size and number of mirabilite crystals increased between −15 and −20°C. For example, a new group of mirabilite crystals is visible below the bubble in the center tube (arrow 4). Few sulfate ions remain in solution at −20°C and there was little additional change in the size or number of mirabilite crystals at −25°C. Further increases in n for the brine also mean less contrast for mirabilite crystals in brine, making them harder to see and reducing their light scattering. No individual hydrohalite crystals were identified in our images. Nevertheless, the brine tube on the left side of Fig. 8.12(c) became entirely opaque (arrow 5) and this is almost certainly due to the precipitation of hydrohalite crystals. Laboratory experiments (Maykut and Light, 1995) suggest that, on the macroscale, hydrohalite and ice frequently form a closely-knit crystalline compound. Unlike mirabilite, which always appeared to remain separate from the ice, the formation of hydrohalite crystals was observed to be closely associated with ice crystal formation. It might be expected that this association would normally lead to brine inclusions being filled with an ice-hydrohalite slush, as appears to be the case in Fig. 8.12(c). However, the formation of this slush was only occasionally observed in tubes and rarely in brine pockets, even at −30°C. Whether hydrohalite crystals tended to nucleate on pocket walls or whether they remained separate from the ice is unclear, but it appears that hydrohalite crystals do not generally exceed 0.01 mm in pockets or tubes.

8.4.4.2 Warming sequence

Fig. 8.13 shows how a group of brine pockets responded as temperature was increased from (a) −13°C to (b) −8°C to (c) −4°C, then decreased to (d) −13°C. Initially the four large inclusions were separate, having an average γ of 1.7 and sizes varying between 0.1 and 0.2 mm in length. Two of the inclusions contained bubbles, but no mirabilite crystals were visible. The predicted increase in brine volume between −13 and −8°C is 45%. The observed increases for the four pockets ranged between 32% and 57%, with the average being 48%. Volume estimates were made by assuming that the inclusions had horizontal aspect ratios of unity and grew the same amount in all horizontal directions. Both the horizontal and vertical dimensions of the pockets were observed to increase. The vertical dimension generally increased slightly faster, resulting in a small increase of γ. Examination of other pockets elsewhere in the same thin section showed similar results. It appears that a slight increase in γ was typical when the pockets became fairly warm. Although existing bubbles increased appropriately in size as the ice warmed, no new bubbles were observed to form in the pockets. Based on the observed increase in brine volume between −13 and −8°C and the difference in brine and ice densities, the pocket at the top of the image, if it were isometric in the horizontal plane,

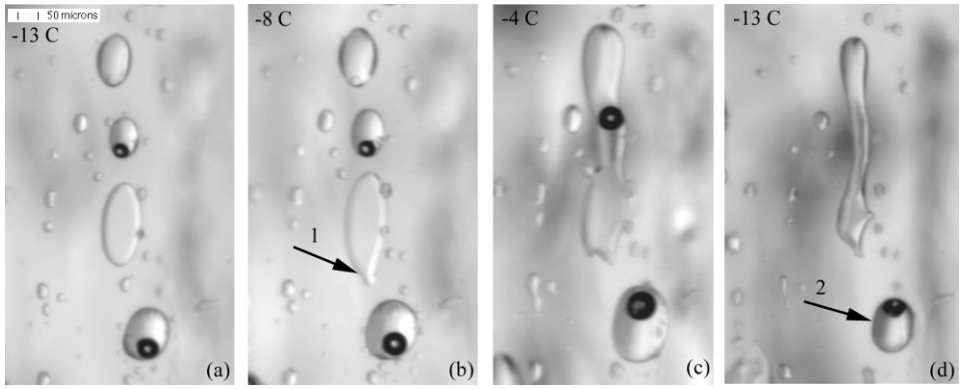


Fig. 8.13. Warming sequence showing changes in ice structure at (a) −13°C, (b) −8°C, and (c) −4°C. Note enlargement and merging of the three brine pockets in the center of the image. The sample was then cooled to (d) −13°C. Arrow 1 indicates merged inclusions, and arrow 2 indicates an isolated pocket with reversible changes. Hysteresis is evident as the large inclusion in (d) remained merged after it had again cooled.

should have formed a bubble with a diameter of approximately 0.028 mm. Such a bubble would have been large enough to be easily resolved and would have been readily apparent. The failure of new bubbles to nucleate in enlarging brine pockets is surprising, but typical of all the scenes examined.

Merging of inclusions was common at high temperature. Note, for example, the third inclusion from the top in Fig. 8.13(b) which combined with a small inclusion below it to produce a single inclusion of irregular shape (arrow 1). When the sample was then warmed to −4°C, three of the four prominent inclusions merged to form a single inclusion with $\gamma = 6$. Because these small inclusions typically form in vertical strings, it was common to observe strings of pockets being transformed to tubes as the ice warmed. This process tends to make the anisotropy of the structure more pronounced at higher temperatures, but it is not immediately clear how this evolution affects the optical properties of the ice.

When the sample was subsequently cooled from −4 to −13°C, the bubble in the newly merged inclusion disappeared, while the merged inclusion remained intact. No new mirabilite crystals were visible. The pocket at the bottom of the image (arrow 2) underwent changes that seem reversible, whereas the other three underwent changes that do not appear to be immediately reversible. Grenfell (1983) suggested that inclusions of large aspect ratio may divide up into smaller inclusions when cooled, but this was not observed in our laboratory samples.

Fig. 8.14 presents a series of lower magnification images that show a broader range of structural changes as the ice was warmed from −15 to −5 to −2°C. At −15°C the inclusions were distinct and well-separated but, as the sample warmed, the inclusions became plump and rounded, and many began to merge. The net effect is that inclusions of brine and gas grew larger, yet decreased in number. These decreased numbers were a direct result of merging, as all inclusions were accounted for. Between −15 and −5°C, brine pocket density decreased from 51 per mm^3 to 46 per mm^3 and bubble density decreased from 5.4 per mm^3 to 4.7

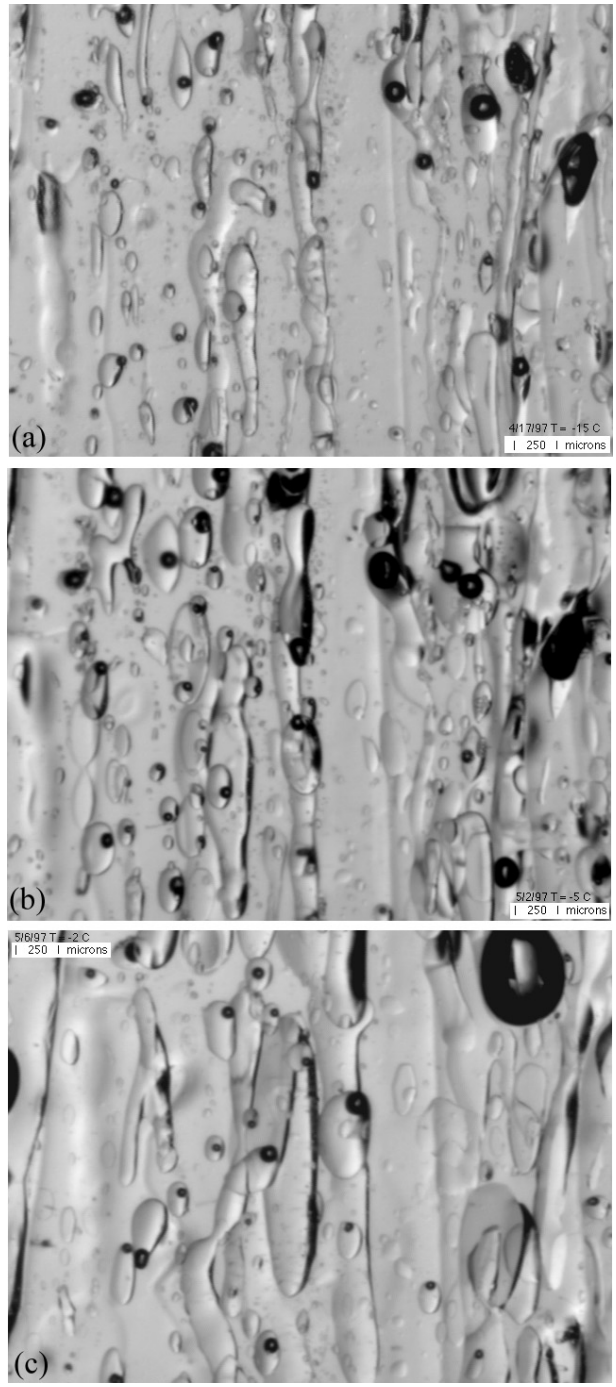


Fig. 8.14. Sequence of images showing changes in the microstructure of first-year congelation ice during warming: (a) −15°C, (b) −5°C, and (c) −2°C. All inclusions enlarged with shapes becoming more plump and rounded. Image width is 3.5 mm.

per mm^3. Between -5 and -2°C, these values further decreased to 35 per mm^3 (69% of the value at -15°C) and 2.7 per mm^3 (50% of the value at -15°C) for brine pockets and bubbles respectively. Number densities in this particular scene are quite large and should not be taken to represent overall averages.

8.4.4.3 Temperature-dependent equivalent cross-sectional area

The temperature dependence of ψ for the various types of scatterers in sea ice was estimated using the freezing equilibrium parameterizations of Cox and Weeks (1983) and guidance from the warming and cooling image sequences. Although brine pockets were observed to become slightly more spherical upon cooling and slightly more elongated upon melting, the overall change in average aspect ratio was small. A model which specifies that melting and freezing take place equally on all inclusion surfaces would produce decreasing aspect ratios for melting and increasing aspect ratios for freezing, contrary to the observations. To estimate $\psi_{\mathrm{bp}}(T)$ for brine pockets, inclusion volumes were allowed to increase as specified by freezing equilibrium but γ was held constant at the value observed at -15°C. No merging between inclusions was considered. This has the effect of shifting the γ curve in Fig. 8.7 to the right as the population warms. Values for r_{eq} and N_{eq} were then computed at various temperatures, treating all brine pockets as prolate ellipsoids. The resulting $\psi_{\mathrm{bp}}(T)$ is shown in Fig. 8.15. It can be seen that, as brine pockets grow larger, the scattering cross-sectional area increases, the increase being most pronounced above -5°C.

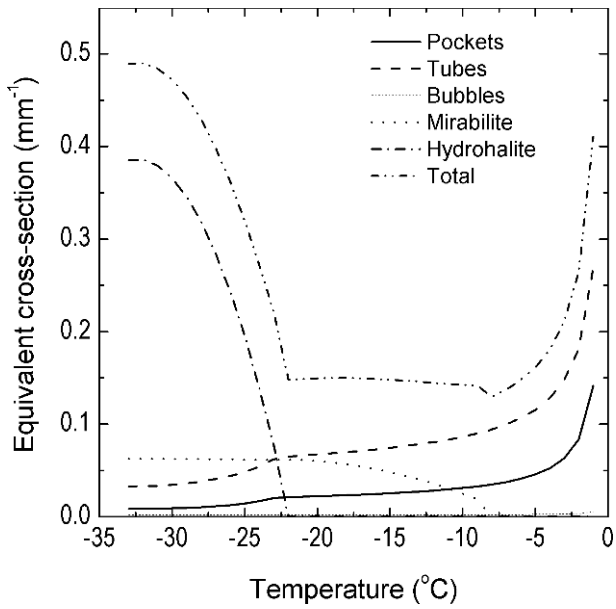


Fig. 8.15. Model results of temperature-dependent equivalent cross-sectional area for brine pockets, tubes, gas bubbles, mirabilite crystals, and hydrohalite crystals.

Tubes showed a different behavior than pockets when the temperature was changed. In general, melting caused tube diameters to increase while tube lengths remained approximately constant. Freezing caused tube diameters to decrease while lengths again remained nearly unchanged. Based on these observations, $\psi_{\mathrm{bt}}(T)$ was calculated by representing the tubes as cylinders and requiring that cylinder length be conserved. The effect of this assumption is that the γ curve in Fig. 8.7 would shift upwards with increasing temperature. The resulting $\psi_{\mathrm{bt}}(T)$ is also shown in Fig. 8.15. The population of tubes visible in the sample clearly has a larger cross-sectional area than the pockets but the calculated values of ψ_{bp} and ψ_{bt} exhibit similar dependence on temperature.

To estimate $\psi_{gb}(T)$ for the bubbles, we first assumed that all bubbles were spherical, that they were located within brine inclusions, and that no new bubbles formed during warming. The change in volume of an existing gas bubble depends on the volume of the brine inclusion in which it resides. Upon warming, more ice melts in a large pocket than in a small one. This means that the volume of a bubble in a large pocket can increase more than if the bubble were located in a smaller pocket. To simplify the problem of determining how r_{eq} depends on temperature, the volume of bubbles observed at $-15°$C was specified to shrink or grow at the bulk rate specified by the freezing-equilibrium relations with fixed bulk density. The resulting $\psi_{gb}(T)$ is shown in Fig. 8.15. The largest values occur at the highest temperatures but, even then, ψ_{gb} does not exceed 3% of the total for the brine inclusions.

These observations were not particularly useful for determining whether precipitated salt crystals tend to grow more in size or number as the ice cools. As a result, we assumed that salt crystals have constant size (0.01 mm edge length) and that changes in temperature cause the crystal number to vary in proportion to changes in the total mass of precipitated mirabilite and hydrohalite. Predicted values of $\psi_m(T)$ and $\psi_h(T)$ for mirabilite and hydrohalite under these assumptions are shown in Fig. 8.15. The maximum ψ_h is almost 6 times larger than ψ_m despite the fact that only 4 times more mass precipitates for the hydrohalite. The stronger effect on ψ_h is due to the fact that the hydrohalite unit cell is closer to equidimensional than the mirabilite unit cell (see Light, 1995). The more equidimensional crystal has higher surface area relative to its volume, and this increases the magnitude of ψ. Had we chosen a smaller crystal size, ψ_m and ψ_h would have been somewhat larger since the total crystal surface area would increase. Of course, the unit cell structure does not uniquely determine the crystal morphology, but lacking specific information on actual crystal dimensions, we used the unit cell as a guide.

8.4.5 Summary of microstructure observations

It can be seen in Fig. 8.15 that equivalent cross-sectional areas for brine pockets, brine tubes, and mirabilite crystals have comparable magnitudes at $-15°$C. As pointed out earlier, tubes contained about 90% of the brine in our samples (see Fig. 8.8), leaving only 10% distributed among pockets. However, Fig. 8.8 also shows that tubes account for only 75% of the total cross-sectional area attributable to brine inclusions. Relative to the tubes, the pockets produce about 3.5 times more equivalent cross-sectional area per unit volume. It is clear from Fig. 8.15 that

tubes provide the largest contribution to ψ of any constituent in the ice at $-15°C$, accounting for roughly half the total in our samples. At $-15°C$, the cross-sectional area for all observed bubbles was an order of magnitude smaller than that for brine pockets, and was only 1% of the total ψ. This is primarily due to the fact that only 5% of all brine inclusions were observed to contain a bubble.

Since the total volume occupied by mirabilite crystals is small (only about 3% of the brine volume at $-15°C$) there was no direct way to determine a size distribution for these crystals, as was done for the brine and gas inclusions. Because the salts precipitate in response to freezing equilibrium, and because their size is generally small compared to the brine inclusions, their presence is tightly tied to the chemical properties of sea ice, and only loosely tied to the inclusion size distributions for tubes, pockets, and gas bubbles. While the observations indicate that mirabilite crystals too large to form in smaller inclusions do sometimes occur in larger brine tubes, most of the precipitated salt crystals do not appear to grow larger than 0.01 mm. In general, we believe the precipitation patterns of crystals to be independent not only of processes occurring at the growth interface, but also of the age and type of the ice. Rather, crystal characteristics are likely to depend primarily on the ice temperature and the chemical composition of the included brine.

Fig. 8.15 shows three distinct temperature regimes: (i) $T < -23°C$, (ii) $-23°C \leq T < -8°C$, and (iii) $T \geq -8°C$. Below $-23°C$, the total ψ is increasingly dominated by the precipitation of hydrohalite crystals. While we do not yet have detailed information on the size and number distributions of these crystals, it is clear from visual observation that their presence causes the ice to become highly scattering at these temperatures (Light, 1995). The cross-sectional area for hydrohalite shown in Fig. 8.15 was predicted using an effective crystal edge length of 0.01 mm at all temperatures. It is still unclear whether hydrohalite crystals tend to nucleate on inclusion walls or within the brine, whether the crystals become incorporated into the ice as the inclusions shrink, or what conditions promote the formation of slush-filled tubes. Furthermore, it is not clear to what extent the brine may become supersaturated with respect to hydrohalite. Regardless of the exact details, the precipitation of hydrohalite is clearly the major factor determining the total cross-sectional area in cold ice, although the role of scattering by individual crystals, in contrast to scattering created by an ice-hydrohalite slush has not been investigated.

The estimated magnitude of ψ is fairly sensitive to our treatment of the salt crystals. Higher values of ψ for the salts would result from smaller, more numerous crystals since the total mass of the precipitate is constrained by the salinity and temperature of the ice. The assumption that the effective crystal edge length is independent of temperature is likely an over-simplification. In fact, the imagery shows both examples of mirabilite crystal growth and additional crystal nucleation as the ice was cooled below $-15°C$ (Fig. 8.12). By assuming a constant crystal size, we preclude the possibility of crystal growth at temperatures below $-15°C$. If we were to permit individual crystals to grow as the ice cools, ψ would increase less rapidly than the calculation shown in Fig. 8.15. However, at temperatures below $-15°C$ the change in ψ_m would be small because few additional sulfate ions are

available to precipitate, regardless of whether new crystals nucleate or existing crystals grow.

At temperatures higher than $-8°C$, $\psi(T)$ is dominated by changes in the geometry of the brine inclusions. Values shown in Fig. 8.15 are based strictly on observed and calculated changes in the size of individual tubes and pockets, without regard to possible merging among neighboring inclusions. A realistic structural-optical model, however, may need to take into account the effects of such merging (Grenfell, 1983). How inclusions enlarge determines how l changes with temperature, but how inclusions merge affects how number densities change with temperature. Merging depends strongly on the spatial distribution of inclusions. Tightly packed inclusions were observed to begin merging at moderate temperatures ($T < -8°C$), while more widely-spaced inclusions did not begin to merge until T reached -2 or $-3°C$. Between -15 and $-5°C$, number density was observed to decrease from $26\,mm^3$ to 20 per mm^3 (23%) for the mosaic in Fig. 8.5. In another scene, number density decreased by 65% between -5 and $-2°C$. Such decreases in number density could reduce the predicted increases in $\psi(T)$.

At intermediate temperatures between -23 and $-8°C$, the change in ψ with temperature is surprisingly weak due to competing effects from the brine inclusions and mirabilite crystals. Warming causes brine inclusions to enlarge and increase their cross-sectional area while, at the same time, mirabilite crystals dissolve and decrease their cross-sectional area. These two effects tend to offset one another. In this regime, the microstructure and the chemistry are of nearly equal importance. Although changes in $\psi(T)$ are sensitive to the assumed crystal size, the estimates shown in Fig. 8.15 assume that the size of the mirabilite crystals is constant at 0.01 mm. Reducing the average crystal size would result in a greater total surface area, a larger cross-sectional areagor the mirabilite, and a steeper change in the total $\psi(T)$. A larger crystal size would reduce the temperature dependence of the total ψ even further. The balance between the competing effects appears to be largely independent of salinity because both effects are predicted to strengthen as the salinity increases. Regardless, in this temperature regime, the mitigating influences of the mirabilite and brine volume yield a $\psi(T)$ with considerably less temperature dependence than that of colder or warmer ice.

These observations of the microstructure were employed in the development of the structural-optical model. We will now describe the observation of apparent optical properties of the ice samples.

8.5 Apparent optical property observations

Inherent optical properties (IOPs) for sea ice (e.g., scattering and absorption coefficients, phase functions) are generally estimated using a plane-parallel radiative transfer model and field data on albedo and transmissivity. This approach becomes problematic, however, when structural or thermal gradients prevent the ice from being accurately characterized by homogeneous scattering properties. To avoid such problems, we made apparent optical property (AOP) measurements (e.g., backscattered and transmitted radiances) on natural ice core samples in a temperature-controlled laboratory cold room. The samples appeared fairly homo-

geneous, suggesting that IOPs could be directly related to the observed microstructure of the ice.

Observations of the AOPs were made on cylindrical core sections 10 cm in diameter and 5 cm thick. Transmitted, backscattered, and side-scattered spectral radiances were recorded at temperatures between −1 and −35°C. Temperature was normally varied in 2–5°C increments, and the ice allowed to equilibrate for at least 24 hours before new measurements were made.

Samples for the optical observations were prepared at −15°C. Cylinders 5 cm in height (H) and 10 cm in diameter were cut from depths of 70–90 cm within the same ice cores used for the microstructural observations described in section 8.4. The top and bottom faces of the cylinder were initially cut with a band saw and then microtomed. The side walls of the core were hand-polished to remove any frost accumulation. Apparent optical property measurements were made by illuminating the top surface of the cylindrical sample with a normally-incident, collimated direct beam of white light (Fig. 8.16). The incident radiation was supplied by a 500-watt incandescent bulb mounted in a parabolic reflector with a glass faceplate, approximately 3 meters horizontally from the sample. By placing the light source as far from the sample as possible within the laboratory, the sample surface presented a small solid angle to the incident light, causing light incident on the sample to be approximately collimated. As the beam was wider than the sample, it illuminated the entire top surface. A mirror set at a 45° angle reflected the light vertically downward onto the top surface of the sample. The sample rested on a 3 mm thick transparent glass plate. Black baffling was used to reduce stray light reaching the sample surfaces.

Measurements of diffuse light backscattered and transmitted by the sample were made using a single fiber optic probe that could be positioned to monitor 5 different radiances: backscattered at 30° ($L_{\alpha 30}$) and 70° ($L_{\alpha 70}$) from zenith, transmitted at 30° (L_{t30}) and 67° (L_{t67}) from nadir, and sidescattered radiance (L_s) normal to the side surface at $H/2$. Transmission measurements were made at 67° instead of 70° to avoid interference from the instrument housing. These particular zenith and nadir angles were selected because they are the ordinate angles for a 4-point Gaussian quadrature (see Eq. (8.6)).

The fiber optic probe was coupled to a Spectron Engineering spectrophotometer (Model SE 590) which recorded light at wavelengths between 400 and 1000 nm with a spectral resolution of about 3 nm. A personal computer was used to control the instrument and record observed spectra. The optical probe was designed to have a narrow field of view so that it could be used to monitor radiance. The radiance probe was fitted with a collimating lens so that the footprint of the sensor would always fit on the sample.

Since absolute radiometric measurements were not necessary, relative radiance measurements were made. A target made of Spectralon®, a material with precisely known spectral albedo ($\alpha_{cal} = 0.99$ at visible wavelengths) was used to both calibrate the probe and monitor the incident irradiance. The albedo of the target is given by

$$\alpha_{\text{target}} = 0.99 = \frac{\int_{2\pi} \tilde{L}(\theta) \cos(\theta)\, d\Omega}{\int_{2\pi} L_{inc}(\theta) \cos(\theta)\, d\Omega}, \tag{8.5}$$

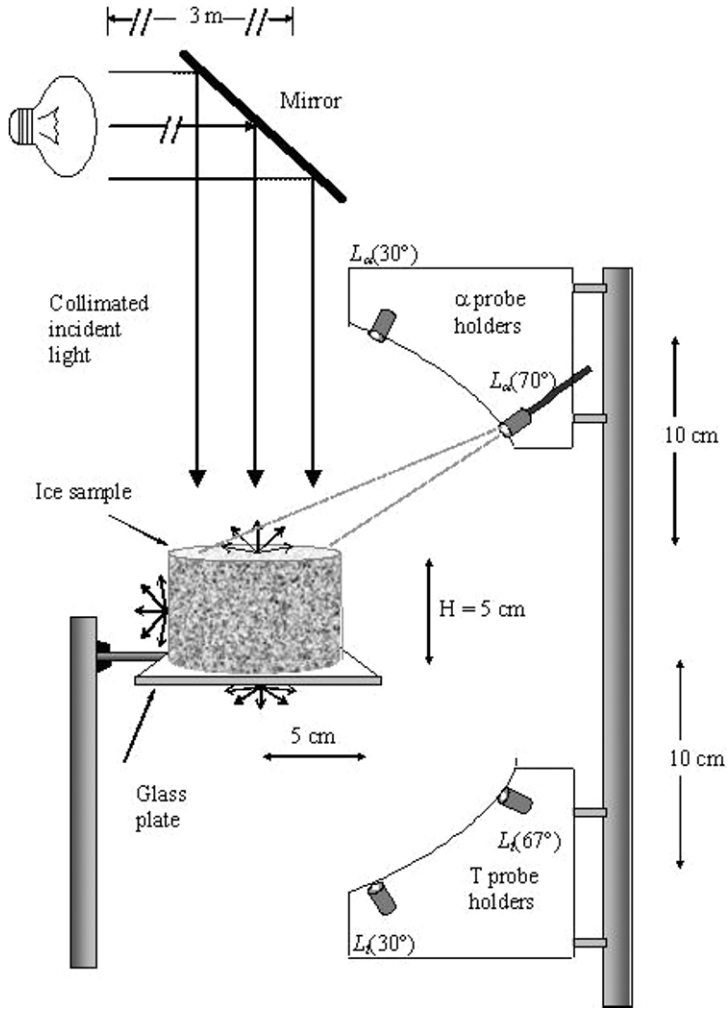


Fig. 8.16. Schematic showing the configuration of incident light, sample, and detector used to collect the optical data.

where $\tilde{L}(\theta)$ is the axially-symmetric radiance field backscattered from the target and $d\Omega$ is the solid angle element over which the radiance field is integrated. Total backscattered irradiance from the target can be estimated from radiance measurements $\tilde{L}'_i$ made at specified angles θ_i using a Gaussian quadrature:

$$\int_{2\pi} \tilde{L}(\theta)\cos(\theta)\,d\Omega \approx 2\pi \sum_i a_i \mu_i x\, \tilde{L}'_i, \tag{8.6}$$

where $\mu_i = \cos\theta_i$, a_i are known quadrature weighting factors (Abramowitz and Stegun, 1965), and x represents the unknown probe throughput which accounts for differences between measured and actual radiances resulting from probe optical characteristics and experiment geometry, i.e. $\tilde{L} = x\tilde{L}'$. Measured radiances L'_i

emanating from the surface of the ice at the quadrature angles can likewise be used to estimate the backscattered irradiance from the ice sample:

$$\int_{2\pi} L(\theta)\cos(\theta)\,d\Omega \approx 2\pi \sum_i a_i \mu_i x L'_i. \tag{8.7}$$

The following expression shows how the relative (normalized by the incident irradiance) radiance $\hat{L}_i$ can be calculated from the measured radiance L'_i, the value of α_{target}, and measurements of backscattered radiance from the target:

$$\hat{L}_i = \frac{xL'_i}{\int_{2\pi} L_{inc}\cos(\theta)\,d\Omega} = L'_i \frac{\alpha_{t\,\mathrm{arg}\,et}}{2\pi \sum_i a_i \mu_i \tilde{L}'_i}. \tag{8.8}$$

Each time optical measurements were made, the calibration target was initially positioned over the sample and $\tilde{L}'_{\alpha 30}$ and $\tilde{L}'_{\alpha 70}$ recorded. The target was then removed to allow measurement of $L'_{\alpha 30}$, $L'_{\alpha 70}$, L'_{t30}, L'_{t67}, and L'_s. Finally, the target was replaced and a second calibration carried out to verify the stability of the source during the measurement sequence. An empirical correction of 4% was made to account for the difference in distance between the probe and the surfaces of the calibration target and ice sample.

Data on spectral radiances were recorded for the ice sample as it was progressively cooled from −15 to −20, −25, and −33°C, then warmed to −15, −11, −7, −5, −4, −2, and −1°C. Typical results are illustrated by the transmission spectra $L_{t30}(\lambda, T)$ shown in Fig. 8.17. Error bars shown on the $T = -15$°C curve indicate uncertainties of ±2%, estimated from the standard deviation of a series of repeated radiance measurements. This variability resulted primarily from random

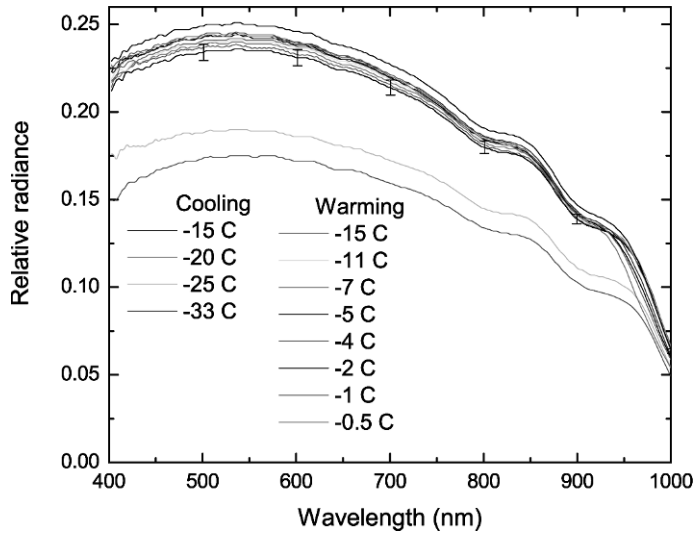


Fig. 8.17. Temperature-dependent changes in spectral radiance transmitted at 30° from nadir. Radiances were measured through the glass plate used to support the sample, then normalized by the incident irradiance. [For the color version of this figure, see the color plate section.]

errors associated with the position of the probe within the probe holders. Normalized $L_{t30}(\lambda, T)$ radiance maxima generally occurred at 530–545 nm and peak values ranged from 0.175 at −33°C to 0.251 at −5°C. Radiances recorded at lower temperatures showed large variability; those recorded at high temperature tended to be more tightly clustered. All transmitted radiances were measured through the glass plate used to support the sample.

After optical measurements were completed, bulk physical properties were determined for the sample. The salinity and density of the optical sample were measured to be 4.7‰ and $0.915 \pm 0.020\,\mathrm{g\,cm^{-3}}$, respectively. These values, used in conjunction with the equations of freezing equilibrium (Cox and Weeks, 1983), gave a brine volume (V_{brine}) of 1.9% and a gas volume (V_g) of 0.95% at −15°C.

The AOP observations were used in conjunction with a radiative transfer model specifically designed to solve transport on a cylindrical domain with refractive top, bottom, and side boundaries. Use of this model was necessary to account for the finite horizontal extent of the laboratory ice samples. In this way, losses out the sidewalls could be accounted for when interpreting the optical observations. The radiative transfer model is briefly described in the following section. A full description of the model can be found in Light et al. (2003b).

8.6 Radiative transfer in a cylindrical domain with refractive boundaries

Multidimensional radiative transfer techniques in the geosciences have been developed to study a variety of problems, including the effects of a spherical shell atmosphere (Collins et al., 1972; Adams and Kattawar, 1978; Herman et al., 1994; Rozanov et al., 2000), effects of broken cloud fields on sunlight (Evans, 1993; Roberti et al., 1994; Gabriel and Evans, 1996; Zuidema and Evans, 1998; Marshak et al., 1998; Roberti and Kummerow, 1999), and specialized problems of light transport in the ocean (e.g., Gordon, 1985; Mobley and Sundman, 2003). The simplest solutions use Monte Carlo techniques and have been used successfully for a variety of problems involving non-plane-parallel geometries. While such models have been built for specialized domains, none have yet been designed to treat radiative transfer in a cylindrical sample of sea ice.

Although Monte Carlo solutions are generally the simplest, a number of more sophisticated solutions have been used to examine radiative transfer in axisymmetric domains. For example, Crosbie and Dougherty (1981) employed an integral solution method to solve the transfer equation in a two-dimensional, radially infinite, axisymmetric domain. Wu and Wu (1997) also applied the integral method to study radiative transfer in a cylindrical medium with finite radial extent. Alcouffe et al. (1997) applied a discrete ordinates method to solve the electrically-neutral particle transport equation in a cylindrical domain. None of these solutions, however, take into account either internal or external refractive boundaries, an important element in the treatment of radiative transfer in sea ice (Grenfell, 1991). Radiative transfer solutions designed for the ocean, on the other hand, usually do account for refraction but only at the ocean–air interface. Since none of the existing models was suitable for our needs, we formulated a new Monte Carlo model that solves the

radiative transfer equation for a finite cylindrical domain with refractive boundaries.

For the task of simulating radiative transfer within the laboratory ice samples, the transport equation is solved using Monte Carlo techniques. In *forward* Monte Carlo models, photons which are emitted from the incident source enter the model domain and travel along trajectories determined by the probabilities of scattering and absorption events. If an interaction is determined to be a scattering event, a scattering angle is calculated by sampling the phase function and the direction of the emergent photon trajectory is updated. This process is continued until the photon exits the domain, is absorbed, or reaches a simulated detector, at which time the next photon is introduced into the domain. Basic radiometric quantities are computed from the fraction of incident photons that reach the location of a specified detector.

When few of the incident photons actually reach the detector, forward models can become very inefficient. Because 2-D problems lack translational invariance, forward Monte Carlo models can fail when the probability that an incident photon will strike the detector becomes very small. For this reason, the *backward* Monte Carlo method was developed by Collins et al. (1972) and Adams and Kattawar (1978). The backward method is based on the principle of reciprocity in radiative transfer theory and requires solution to the *adjoint*, or time-reversed problem. In the adjoint problem, photons are generated at the detector and their trajectories traced back to the original source. Photons are tallied as they leave the domain and every photon generated contributes to the solution. The backward Monte Carlo method has been implemented for radiative transfer problems in the ocean (Gordon, 1985; Mobley and Sundman, 2003), within clouds (Evans, 1993; Roberti et al., 1994), and for limb viewing radiances in the atmosphere (Oikarinen, 2001). We next discuss the development of the 2-D radiative transfer model and demonstrate how it can be used in the analysis of optical data.

8.6.1 Model overview

For this task, we chose to develop a backward, axisymmetric Monte Carlo model with multiple refractive layers. We closely follow the treatment described by Gordon (1985) for radiative transfer in the ocean and generalize it to accommodate multiple refractive boundaries in a cylindrical domain.

We consider a right circular cylinder whose axis is oriented vertically and whose upper surface is illuminated with either diffuse or normally incident collimated radiation. Vertical variations in inherent optical properties are simulated using horizontal layers, each with specified vertical extent (h_n) and uniform properties. Similarly, radial variations are simulated using concentric cylindrical shells with radii ρ_m (see Fig. 8.18). IOPs, including the real refractive index (n), absorption coefficient (κ), scattering coefficient (σ), and scattering phase function ($p[\boldsymbol{\xi}' \rightarrow \boldsymbol{\xi}]$), must be specified for each layer of each shell.

The equation governing the transfer of radiation can be written:

$$(\boldsymbol{\xi} \cdot \nabla) L(\boldsymbol{r}, \boldsymbol{\xi}) + k(\boldsymbol{r}) L(\boldsymbol{r}, \boldsymbol{\xi}) = \sigma(\boldsymbol{r}) \int_{4\pi} p(\boldsymbol{r}, \boldsymbol{\xi}' \rightarrow \boldsymbol{\xi}) L(\boldsymbol{r}, \boldsymbol{\xi}')\, d\Omega(\boldsymbol{\xi}') + Q(\boldsymbol{r}, \boldsymbol{\xi}), \quad (8.9)$$

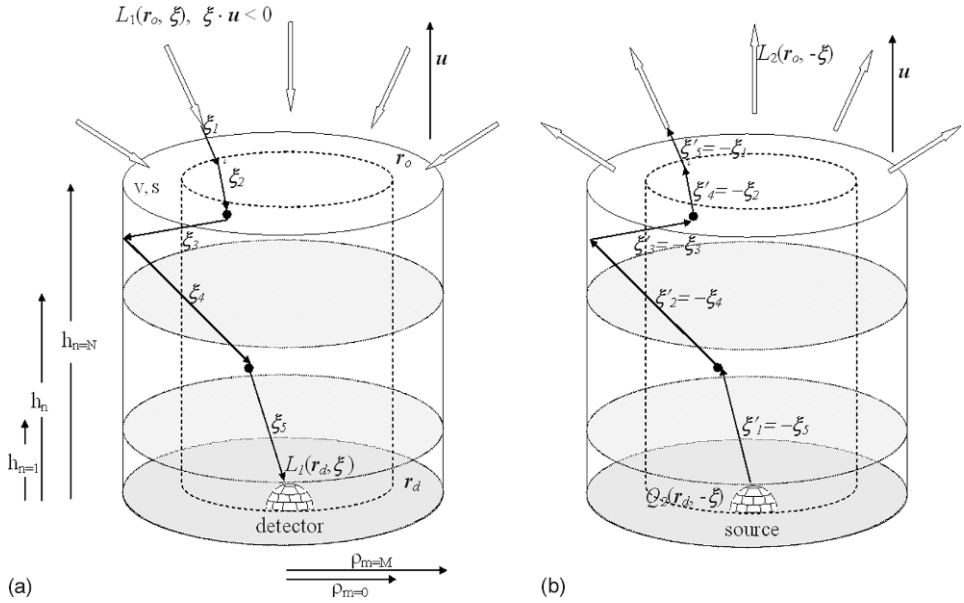


Fig. 8.18. Schematic of the (a) forward and (b) adjoint problems for diffuse incident radiation.

where the radiance $L(\boldsymbol{r}, \boldsymbol{\xi})$ is specified at vector location $\boldsymbol{r}$ in the direction of unit vector $\boldsymbol{\xi}$ ($\boldsymbol{\xi} > 0$ in downward direction) and the integral over solid angle Ω is evaluated over 4π steradians. The beam attenuation coefficient, $k(\boldsymbol{r})$, is the sum of $\kappa(\boldsymbol{r})$ and $\sigma(\boldsymbol{r})$. $Q(\boldsymbol{r}, \boldsymbol{\xi})$ is the intensity (power per unit solid angle) per unit volume of all sources at $\boldsymbol{r}$ in direction $\boldsymbol{\xi}$. The phase function $p(\boldsymbol{r}, \boldsymbol{\xi}' \rightarrow \boldsymbol{\xi})$ describes how light with direction $\boldsymbol{\xi}'$ is scattered into direction $\boldsymbol{\xi}$ at location $\boldsymbol{r}$. The model can treat any phase function. For the purpose of illustrating this model, we use a Henyey–Greenstein analytic function (Henyey and Greenstein, 1941) because of its convenience and efficiency. The phase function is assumed to have azimuthal symmetry about the direction of the incident photon, an appropriate assumption for scattering by equivalent spheres.

The forward problem for diffuse illumination is illustrated in Fig. 8.18(a). The incident radiance is $L_1(\boldsymbol{r}_o, \boldsymbol{\xi})$, where $\boldsymbol{r}_o$ refers to the top surface of the cylinder and subscript 1 refers to the forward problem. The vector normal to the top surface is denoted $\boldsymbol{u}$. Photons refract, reflect, and scatter as they enter and propagate through the medium. The figure shows a series of plausible trajectories for a photon as it transits through the domain to a transmissivity detector. The radiance at the location of the detector ($\boldsymbol{r}_d$) is denoted $L_1(\boldsymbol{r}_d, \boldsymbol{\xi})$. The corresponding downwelling irradiance at the detector ($E_d[\boldsymbol{r}_d]$), normalized by the incident downwelling diffuse irradiance ($E_d[\boldsymbol{r}_o]$), is:

$$\frac{E_d(\boldsymbol{r}_d)}{E_d(\boldsymbol{r}_o)} = \frac{-\int_{\boldsymbol{\xi}\cdot\boldsymbol{u}<0} L_1(\boldsymbol{r}_d, \boldsymbol{\xi})\boldsymbol{\xi}\cdot\boldsymbol{u}\, d\Omega}{-\int_{\boldsymbol{\xi}\cdot\boldsymbol{u}<0} L_1(\boldsymbol{r}_o, \boldsymbol{\xi})\boldsymbol{\xi}\cdot\boldsymbol{u}\, d\Omega}. \tag{8.10}$$

The $\boldsymbol{\xi} \cdot \boldsymbol{u} < 0$ bound on the solid angle integral indicates that the integration is to be carried out only for downwelling directions. We will frequently refer to the ratio $E_d(r_d)/E_d(r_o)$ in this study.

The adjoint problem is illustrated in Fig. 8.18(b). Case and Zweifel (1967) demonstrated that Eq. (8.9) possesses unique solutions for a closed convex volume, V, bounded by surface, S, with single scattering albedo $\varpi_o \equiv \sigma(\boldsymbol{r})/k(\boldsymbol{r}) < 1$, provided the incident radiance, IOPs, and sources within V are known. If the phase function is time-reversal invariant, i.e. if $p(\boldsymbol{r}, \boldsymbol{\xi}' \rightarrow \boldsymbol{\xi}) = p(\boldsymbol{r}, -\boldsymbol{\xi} \rightarrow -\boldsymbol{\xi}')$, the following reciprocity relationship relates the forward (subscript 1) and adjoint (subscript 2) problems (Case, 1957):

$$\int_{\xi \cdot u < 0} d\Omega \int_S dS \, |\xi \cdot u| \; [L_1(\boldsymbol{r}_o, \xi) L_2(\boldsymbol{r}_o, -\xi) - L_2(\boldsymbol{r}_o, \xi) L_1(\boldsymbol{r}_o, -\xi)]$$
$$= \int_{4\pi} d\Omega \int_V dV [L_1(\boldsymbol{r}, -\xi) Q_2(\boldsymbol{r}, \xi) - L_2(\boldsymbol{r}, \xi) Q_1(\boldsymbol{r}, -\xi)], \tag{8.11}$$

where $L_1(\boldsymbol{r}_o, -\boldsymbol{\xi})$ and $L_2(\boldsymbol{r}_o, -\boldsymbol{\xi})$ are the radiances leaving the domain at the top surface in the forward and adjoint problems, respectively. Since there are no internal sources in the forward problem considered here, $Q_1(\boldsymbol{r}, -\boldsymbol{\xi}) = 0$; and, equivalently, since there is no incident radiation in the adjoint problem, $L_2(\boldsymbol{r}_o, \boldsymbol{\xi}) = 0$, Eq. (8.11) becomes

$$\int_{\xi \cdot u < 0} d\Omega \int_S dS \, |\xi \cdot u| \; L_1(\boldsymbol{r}_o, \xi) L_2(\boldsymbol{r}_o, -\xi)$$
$$= \int_{4\pi} d\Omega \int_V dV [L_1(\boldsymbol{r}, -\xi) Q_2(\boldsymbol{r}, \xi)]. \tag{8.12}$$

The normalized, transmitted irradiance in Eq. (8.10) can be computed from Eq. (8.12) by choosing a suitable value for $Q_2(\boldsymbol{r}, \boldsymbol{\xi})$. $Q_2(\boldsymbol{r}, \boldsymbol{\xi})$ is the function governing the angular distribution with which photons are released at the detector location in the adjoint problem. To simulate a transmitted irradiance detector, we choose

$$Q_2(r, \xi) = \begin{bmatrix} (\xi \cdot u) \delta(r - r_d) \mathrm{J}_o & \xi \cdot u > 0 \\ 0 & \xi \cdot u < 0 \end{bmatrix}, \tag{8.13}$$

where J_o is the total number of photons released. For a transmitted irradiance detector, Eq. (8.13) specifies that photons be released only from the upper hemisphere of the detector with a $\boldsymbol{\xi} \cdot \boldsymbol{u}$ dependence. This is simply a mathematical expression of how a cosine detector *collects* photons.

If we assume that the incident radiance is isotropic with value L_o, then the incident irradiance is $E_d(\boldsymbol{r}_o) = \pi L_o$. Substituting Eq. (8.13) into Eq. (8.12) and evaluating the RHS of Eq. (8.12) for the upward-looking hemisphere allows us to rewrite Eq. (8.10) as

$$\frac{E_d(\boldsymbol{r}_d)}{E_d(\boldsymbol{r}_o)} = \frac{\int_S dS \int_{\xi \cdot u > 0} d\Omega \, |\xi \cdot u| \, L_2(\boldsymbol{r}_o, -\xi)}{\pi J_o}. \tag{8.14}$$

This equation can now be evaluated using a Monte Carlo simulation. Photons simulating Q_2 are released at $\boldsymbol{r}_d$ and are followed through the domain until they

exit the top surface as $L_2(\boldsymbol{r}_o, -\boldsymbol{\xi})$. The numerator is simply the total intensity of photons that leave the top surface of the domain. All photons that pass through $\boldsymbol{r}_o$ constitute the radiance $L_2(\boldsymbol{r}_o, -\boldsymbol{\xi})$ and contribute to the solution integral. These results allow irradiances and radiances to be computed at the location and direction of any detector simply by computing the power associated with photons leaving the top of the cylinder and comparing it with the power associated with photons leaving the adjoint source.

8.6.2 Implementation

8.6.2.1 Photon release

It is convenient to think of each photon released at the location of the detector (henceforth called the 'source', see Fig. 8.18(b)) as representing a *packet* of photons. The number of photons in the packet, or the intensity of the associated beam, is characterized by a statistical weight W. Initially, $W = 1$. By appropriately scaling W, the number of photons in a packet can be modified to reflect the portion of the packet surviving absorption or interaction with a refractive boundary. The initial position of each packet is $\boldsymbol{r}_d$, and the direction of the packet trajectory is denoted by polar (θ) and azimuthal (ϕ) angles. The source is a hemisphere divided into a specified number of equal solid-angle bins. Initial values of (θ, ϕ) are assigned the central polar and azimuth angles of the bin from which the packet is released. For an irradiance source, all bins in the hemisphere are active and equal numbers of photons are released from each. For a radiance source, only bins within the acceptance angle of the viewing direction are active.

8.6.2.2 Distance between collisions

The distance (d) a packet travels in direction $\boldsymbol{\xi}$ before it undergoes a scattering or absorption event is determined by $k(\boldsymbol{r})$. Because the packet trajectory is determined one scattering event at a time, the intensity is attenuated with distance exponentially for each segment. A distance to the next interaction is selected through a process of choosing a random number and using it to sample a cumulative distribution function. This is generally known as the *fundamental principle* of Monte Carlo simulation (Cashwell and Everett, 1959) and is used extensively in the present model. A random number R selected on the interval $0 \leq R < 1$ is used to sample the cumulative distribution corresponding to an exponential probability distribution function, and

$$d = -\frac{1}{k(\boldsymbol{r})} \ln[1 - R]. \tag{8.15}$$

Since k represents the total probability of all scattering and absorption processes, this calculation gives the distance to the next interaction, regardless of whether it is a scattering or absorption event. The average over many computations of d is the mean free path and is equal to $1/k(\boldsymbol{r})$.

8.6.2.3 Photon interactions

Each time the distance to a collision is computed, it is compared with the distance to the nearest boundary along the direction of travel. The packet is advanced to the location of the nearest interaction, either: (i) an internal horizontal or vertical boundary, (ii) an external horizontal or vertical boundary, or (iii) a scattering or absorption event. If the photon interacts with an internal refractive layer boundary, a random number is drawn and the fundamental principle of Monte Carlo simulation is invoked to determine whether the packet is reflected or refracted. If the packet is reflected, it remains within the same layer with its direction reflected about the normal to the interface. The remaining distance to the collision is computed and compared with distances to boundaries along the new trajectory. If the packet is refracted, it is transmitted to the neighboring layer, where the remaining distance to the collision d' is adjusted to account for differences in the k values between the previous layer and the new layer:

$$d' = d \cdot k_{previous}/k_{new}. \tag{8.16}$$

The statistical weights obey the Fresnel equations whenever a photon interacts with a refractive boundary. As packet weights represent radiances along the direction $\boldsymbol{\xi}$, W for photon packets crossing a boundary with relative refractive index (m) must be modified to conserve energy:

$$W' = W\frac{1}{m^2} \cdot t(\xi,\ m) \tag{8.17}$$

where $t(\boldsymbol{\xi}, m)$ is the Fresnel transmission coefficient at the boundary (Wyatt, 1978; Mobley, 1994).

If the photon interacts with the top of the cylinder, the transmitted portion of the packet contributes to the final solution. Interaction with any refractive external boundary causes the packet weight to be reduced by the reflection coefficient and its trajectory to be reflected by the surface. In this way, packets are forced to remain within the domain until exiting the top surface. Packets that experience total internal reflection continue to trace out trajectories within the medium. To reflect a packet at the curved side wall, the reflection coefficient is calculated using Fresnel's principle relative to the local tangent plane.

If the distance to the collision is less than the distance to any boundary in the direction of the trajectory, the packet will experience a scattering or absorption event. Losing packets to absorption events is computationally inefficient, as terminated packets no longer contribute to the solution. To eliminate this inefficiency, all interactions are forced to be scattering events, and the bias is taken into account by multiplying W by the probability the interaction is a scattering event $(\varpi_o[\boldsymbol{\eta}])$ each time an interaction occurs. By doing this, the packet is weighted by the probability that it has survived this particular interaction. Once a scattering event occurs, the direction of the next trajectory is selected by using the fundamental principle of Monte Carlo simulation to sample the polar and azimuthal angles of the phase function. The photon position remains unchanged, but θ and ϕ are reassigned.

8.6.2.4 Photon counting

When a packet hits the top boundary of a non-refractive medium, it will leave the domain, contributing its weight, W, to $\int L_2(r_o, -\xi)$ in the numerator of Eq. (8.14). When a packet hits the top boundary of a refractive medium, a portion of it will be reflected and a portion will be transmitted, unless it suffers total internal reflection. In the case of partial transmission, the transmitted portion contributes to the solution and the reflected portion remains within the medium. It continues to scatter until its weight drops below a prescribed threshold (typically $W = 10^{-6}$), potentially interacting with the top surface numerous times, each time contributing to the solution integral.

Equation (8.14) is straightforward to apply when the incident radiation is isotropic. When the incident radiation has an angular dependence, $\mathrm{C}(\boldsymbol{\xi})$, exiting photons must be weighted by $\mathrm{C}(\boldsymbol{\xi})$ as they contribute to $L_2(\boldsymbol{r}_o, -\boldsymbol{\xi})$. When the incident radiation is collimated within some solid angle $\Delta\Omega(\boldsymbol{\xi}_o)$, $\mathrm{C}(\boldsymbol{\xi}) = \delta(\Delta\Omega(\boldsymbol{\xi}_o))$ and only photons exiting within $\Delta\Omega(\boldsymbol{\xi}_o)$ have non-zero weighting. This realization is necessary to simulate the AOP observations, but leads to very inefficient, high variance solutions as $\Delta\Omega(\boldsymbol{\xi}_o)$ decreases. Gordon (1985) overcame this problem by reformulating the solution for collimated incident irradiance by rederiving Eq. (8.14) and separating the radiance field at the detector into two components to be treated separately: radiation scattered within the domain and radiation falling on the detector without having been scattered. Fig. 8.19(a) illustrates the forward problem for the case of collimated incidence. The incident beam is denoted $L_1(\boldsymbol{r}_o, \boldsymbol{\xi}_o)$, and the transmission detector records both scattered and unscattered light. The scattered light results from photons that interact with the medium along trajectories that eventually intersect the detector. The unscattered portion ($L_1[\boldsymbol{r}_d, \boldsymbol{\xi}_o]$) is given by the attenuated incident beam persisting at the location of the detector.

Fig. 8.19(b) illustrates the treatment of the scattered component in the adjoint problem. As in the diffuse case, photons are released at $\boldsymbol{r}_d$ using the distribution function $Q_2(\boldsymbol{r}, \boldsymbol{\xi})$, and their trajectories are traced through the domain. Instead of waiting for photons to exit the top of the domain at angle $\boldsymbol{\xi}_o$, which very few would do, they contribute to the solution integral at each scattering event. At each event, the probability that the photon was scattered out of the direct beam into the current trajectory is recorded.

In the case of a non-refractive domain (e.g., a cloud), or a domain with a single refractive boundary (e.g., the ocean surface), the only unscattered beam is downwelling $[L(\boldsymbol{r}, \boldsymbol{\xi}_o)]$, and is simply the incident collimated beam with intensity attenuated to the depth of the event and adjusted for transmission at the refractive boundary. In the case of a domain with multiple refractive boundaries, the direct beam can have both upwelling and downwelling components at any location above the bottom boundary, and the probability of scattering this beam into the photon trajectory must be included for both directions of direct beam propagation. Each term in the solution sum is the product of the packet weight and the probability that the photon came into that trajectory either from the downwelling or the upwelling portion of the as-yet unscattered beam. The sum of these terms is:

$$\sum_{ij} W_i P_{ij}, \tag{8.18}$$

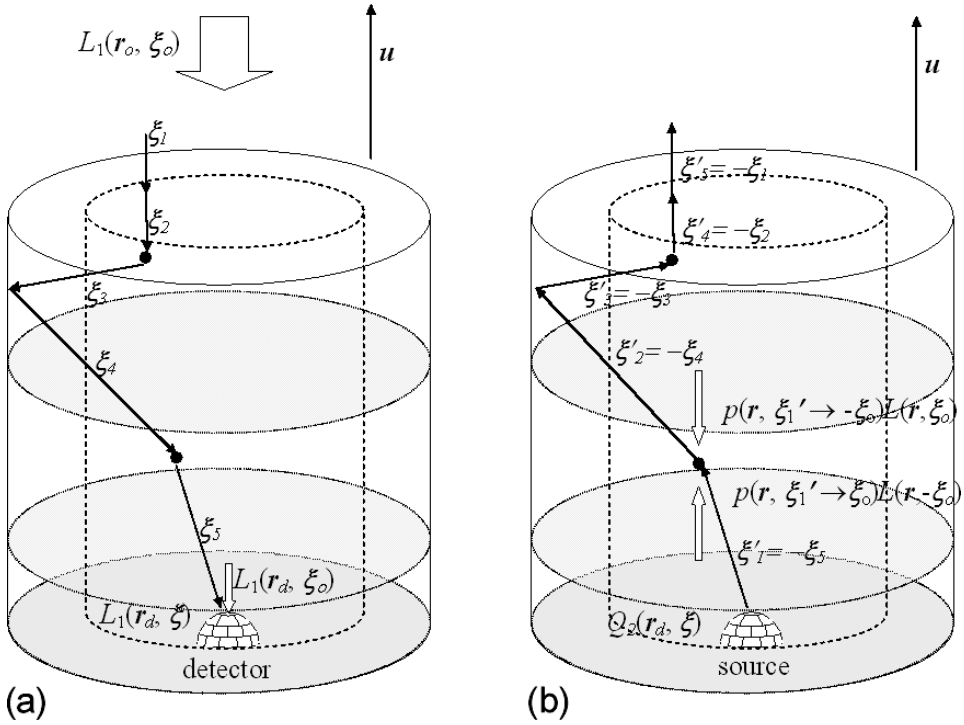


Fig. 8.19. Same as Fig. 8.18, except for direct incident radiation.

where

$$P_{ij} = \varpi_o(\boldsymbol{r})\Delta\Omega(\xi_o) \cdot \{p(\boldsymbol{r}, \xi' \rightarrow -\xi_o)L(\boldsymbol{r}, \xi_o) + p(\boldsymbol{r}, \xi' \rightarrow \xi_o)L(\boldsymbol{r}, -\xi_o)\} \quad (8.19)$$

and $L(\boldsymbol{r}, \boldsymbol{\xi}_o)$ and $L(\boldsymbol{r}, -\boldsymbol{\xi}_o)$ represent the downwelling and upwelling unscattered incident beams at the location of the collision, as shown in Fig. 8.19(b). W_i is the statistical weight of photon i.

Combining the contributions of interacting and non-interacting photons, the normalized downwelling irradiance Eq. (8.14) becomes

$$\frac{E_d(\boldsymbol{r}_d)}{E_d(\boldsymbol{r}_o)} = \frac{\sum_{ij} W_i P_{ij}}{\pi J_o \, |\xi_o u| \, \Delta\Omega(\xi_o)} + \frac{L(\boldsymbol{r}_d, \xi_o)}{L_o}. \quad (8.20)$$

As with diffuse incidence, photon packets are released at $\boldsymbol{r}_d$ with $W = 1$ and allowed to generate trajectories through the medium. The unscattered term is the fraction of the incident beam that reaches the detector without having undergone a scattering event. For a transmissivity detector, this is $L(\boldsymbol{r}_d, \boldsymbol{\xi}_o)$; for an albedo detector, $L(\boldsymbol{r}_d = \boldsymbol{r}_o, -\boldsymbol{\xi}_o)$. Since $\sum_{ij} W_i P_{ij}$ is directly proportional to $\Delta\Omega(\boldsymbol{\xi}_o)$, $E_d(\boldsymbol{r}_d)/E_d(\boldsymbol{r}_o)$ is independent of $\Delta\Omega(\boldsymbol{\xi}_o)$.

8.6.3 Similarity

The optical properties of sea ice are largely a result of absorption incurred by the pure ice matrix plus impurities (Grenfell, 1983; Light et al., 1998) and scattering due to embedded inclusions of brine, gas, and precipitated salts. The absorption of pure ice is well known, but scattering due to inclusions has not been well documented. To specify total scattering, both the scattering coefficient and the phase function must be quantified. Simultaneous inference of these two unknowns can be problematic. In radiative transfer theory, similarity relations sometimes permit reduction in the number of independent IOP variables by one. Here, we will make use of a *similarity parameter*, S, defined by van de Hulst (1980) as

$$S = [1 + (\sigma/\kappa)(1 - g)]^{-1/2}, \tag{8.21}$$

or, equivalently, as

$$S = [(1 - \varpi_o)/(1 - g\varpi_o)]^{1/2},$$

S varies from 0 to 1, approaching 1 as κ or g become large.

Since κ_λ of pure ice at visible wavelengths is well established (Grenfell and Perovich, 1981), and since we understand the total scattering in sea ice to be approximately independent of wavelength for visible light, we concern ourselves only with the magnitude of σ, not σ/κ_λ. If κ_λ is specified, a more simple form of the similarity parameter can be expressed as

$$s = \sigma(1 - g). \tag{8.22}$$

This statement gives a rule by which smaller σ values may be used to compensate for smaller g values. As with S, two domains with the same absorption and identical s values will have nearly identical AOPs [van de Hulst, 1980].

An additional benefit of using a similarity relation is to reduce the computation time for forward-peaked phase functions (Graaff et al., 1993). Solution time for the Monte Carlo model is a serious concern and depends strongly on the inherent optical properties of the medium. While domains with large optical depth take longer to execute, the computation time for any domain depends strongly on the asymmetry parameter. The more forward-peaked the phase function, the longer it takes to obtain a solution. Experiments with the model have demonstrated that there exists a direct relationship between g and computation time. As g increases from 0.9 to 0.95, computations take approximately an order of magnitude longer; an increase from 0.9 to 0.98 takes approximately two orders of magnitude longer. The similarity parameter can thus aid in the calculation of radiative transfer when the phase function is extremely forward peaked.

8.6.4 Simulation of laboratory observations

To simulate the observations reported in section 8.5, the ice sample was modeled as a refractive cylinder with spatially uniform σ, κ_λ, and g, while the supporting glass plate was modeled as a non-scattering second layer with a refractive index of 1.50 and a spectral absorption coefficient that was estimated with a simple transmittance measurement in the laboratory. It was initially assumed that scattering in

the ice could be described by a Henyey–Greenstein phase function with $g = 0.95$. The Monte Carlo Model (MCM) was used to calculate $L_{\alpha 30}$, $L_{\alpha 70}$, L_{t30}, L_{t67}, and L_s for a wide range of s values. Observed radiances were then compared with the predicted radiances, and the value of s that best explained all five observed spectra was selected at each temperature. Derived values for $s(T)$ are shown in Fig. 8.20. The cooling sequence is indicated by open symbols and the warming sequence by solid symbols. Both sequences were initiated at −15°C. Symbol numbers indicate the order in which measurements were made. The overall range of s varied from $8.6\,\mathrm{m}^{-1}$ at −5°C to $15.6\,\mathrm{m}^{-1}$ at −33°C. This is almost a twofold change in attenuation. Uncertainties are estimated to be approximately the symbol size.

Three distinct temperature regimes can be identified in Fig. 8.20.

(a) $T < -23$°C: Values of s increased by nearly 70% as the sample was cooled between −20 and −33°C. Such strong increases in attenuation by scattering were almost certainly the direct result of the precipitation of hydrohalite crystals which should begin to occur at −23°C.
(b) -23°C $< T < -8$°C: Attenuation due to scattering at intermediate temperatures appeared to be almost constant. Between −15 (point 5) and −5°C (point 8), for example, s changed by less than 5%. Two competing factors appear to contribute to this behavior. As sea ice warms across this temperature regime, there are moderate increases in scattering due to increased brine and gas volume. At the same time, there are decreases in scattering due to the decreasing mass of precipitated mirabilite crystals. To a large extent, these effects appear

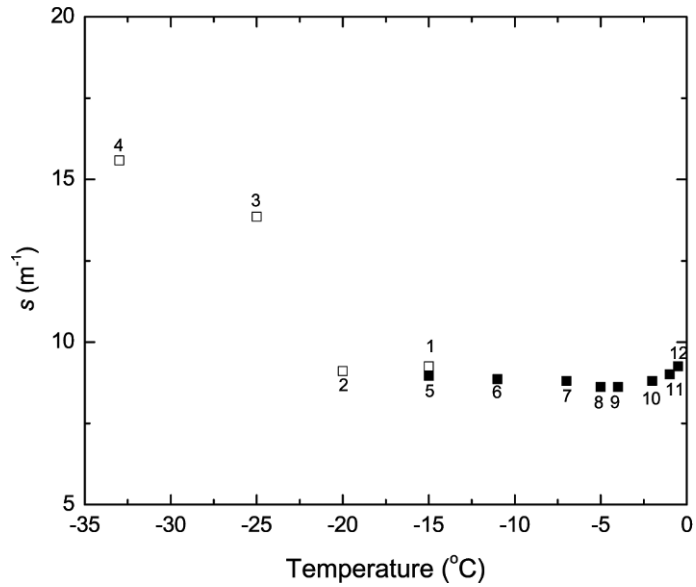


Fig. 8.20. Derived values of the similarity parameter s as a function of temperature. Values were derived from observed radiances by comparison with Monte Carlo model calculations.

to counteract one another. Scattering by mirabilite crystals eventually ceases above −8°C, where all mirabilite crystals should completely dissolve.

(c) $T > -8$°C: At temperatures above the mirabilite precipitation (points 7–12), attenuation due to scattering decreased slightly as the sample was warmed to −4°C (point 9), then increased monotonically as the temperature approached −1°C (point 12). The total increase in s was less than 10% between −8°C and −1°C, although the ice experienced an eightfold increase in brine volume and an almost twofold increase in gas volume. Such large increases in brine and gas volume were expected to produce correspondingly large increases in scattering, but the observed optical changes were surprisingly small. The effort to attribute these changes in scattering quantitatively is now described in the development of the structural-optical model.

8.7 Structural-optical model

The structural-optical model is based on the theoretical framework of Grenfell (1983, 1991). The model uses direct observations of inclusion number densities and size distributions obtained from microstructural imagery to predict $\kappa(T, \lambda)$, $\sigma(T)$, and $g(T)$ for brine inclusions, gas bubbles, and precipitated salt crystals. Fundamental physical relationships between the ice structure and its optical properties are deduced by iteratively comparing values of s derived from AOP data with values of s predicted with the structural-optical model. These comparisons were used to refine and improve the model.

8.7.1 Structural-optical relationships

In this section we develop explicit relationships between the microstructure of the ice and the inherent optical properties of each type of included scatterer. We also describe how these relationships are used to predict values of κ, σ, and g for the entire sample. These relationships constitute the structural-optical model.

8.7.1.1 Absorption coefficient, κ

The total spectral absorption coefficient, $\kappa_{total}(\lambda)$, can be written as the volume-weighted sum of the individual absorption coefficients ($\kappa_j[\lambda]$) of pure ice, included brine, precipitated salts, and included particulates (Grenfell, 1991):

$$\kappa_{total}(\lambda) = \kappa_{ice}(\lambda)V_{ice} + \kappa_{brine}(\lambda)V_{brine} + \kappa_{ps}(\lambda)V_{ps} + \kappa_{particulates}(\lambda)V_{particulates}. \tag{8.23}$$

where V_j is fractional volume. Because the samples used in this study did not generally contain significant quantities of absorbing particulates, $V_{particulates}$ was assumed to be zero. Since information on absorption by the precipitated salts was not available and since V_{ps} is very small relative to V_{ice} and V_{brine}, effects of κ_{ps} on total absorption were also assumed to be negligible, hence

$$\kappa_{total}(\lambda) \approx \kappa_{ice}(\lambda)V_{ice} + \kappa_{brine}(\lambda)V_{brine}. \tag{8.24}$$

Because V_{ice} is usually much larger than V_{brine} and because $\kappa_{ice}(\lambda)$ is close to $\kappa_{brine}(\lambda)$, $\kappa_{total}(\lambda)$ is approximately equal to $\kappa_{ice}(\lambda)$. $\kappa_{ice}(\lambda)$ is strongly wavelength-dependent, with values less than $0.05\,\mathrm{m}^{-1}$ at blue wavelengths, increasing to more than $10\,\mathrm{m}^{-1}$ in the near-infrared.

8.7.1.2 Scattering coefficient, σ

The temperature-dependent scattering coefficient for each type of inclusion is given by:

$$\sigma(T) = \int_{r_{\min}}^{r_{\max}} Q_{sca}(r_{eq})\pi r_{eq}^2(T)N(r_{eq},T)\,dr_{eq} \tag{8.25}$$

where $Q_{sca}(r_{eq})$ is the scattering efficiency. All the inclusions considered here (i.e., brine, gas, and salt crystals) have sizes much larger than the wavelength, so that Q_{sca} is approximately equal to 2, and $\sigma(T) \approx 2 \cdot \psi(T)$ (See section 8.4.3.2). The calculation of σ(T) for each type of inclusion is described below.

8.7.1.3 Brine inclusions

The total value of ψ_{brine} for all brine inclusions observed in the thin section imagery at $-15°$C was estimated to be $110\,\mathrm{m}^{-1}$, with $30\,\mathrm{m}^{-1}$ being contributed by the pockets (ψ_{bp}) and $80\,\mathrm{m}^{-1}$ by the tubes (ψ_{bt}). However, the brine volume observed in the microstructure sample ($V_{brine} = 1.2\%$) was only 60% of that calculated for the optical sample at $-15°$C. The most likely explanation is that the thin section did not contain many of the larger and more widely dispersed brine tubes present in the core. Such features would, of course, lead to greater scattering in the optical sample. To estimate the effect on ψ_{bt}, we assumed that these larger tubes followed the same size distribution given by Eq. (8.1). We found that extending $N_{brine}(l)$ from $l = 8\,\mathrm{mm}$ to include tubes up to 14.6 mm in length produced a brine volume of 1.9%, in agreement with the optical sample. This increased $\psi_{\mathrm{bt}}(-15°\mathrm{C})$ from 80 to $110\,\mathrm{m}^{-1}$, and yielded estimates for $\sigma_{\mathrm{bp}}(-15°\mathrm{C})$ of $60\,\mathrm{m}^{-1}$ and $\sigma_{\mathrm{bt}}(-15°\mathrm{C})$ of $220\,\mathrm{m}^{-1}$.

Temperature-dependent volume changes for each brine inclusion were predicted from the equations of freezing equilibrium (Cox and Weeks, 1983). Volume increases were assumed to be proportional to the initial size of the inclusion. Pockets were assumed to grow and shrink in length and diameter, retaining constant aspect ratio; tubes were assumed to grow and shrink in diameter and aspect ratio, retaining constant length. These rules constitute our structural model for brine inclusions. Application of this model to the scattering coefficients obtained at $-15°$C allow calculation of $\sigma_{\mathrm{bp}}(T)$ and $\sigma_{\mathrm{bt}}(T)$. The results (Fig. 8.21) indicate that scattering by brine tubes in interior, first-year ice is roughly 4–5 times that of brine pockets at all temperatures. As temperature increases, so does the contribution of brine inclusions to the total scattering coefficient.

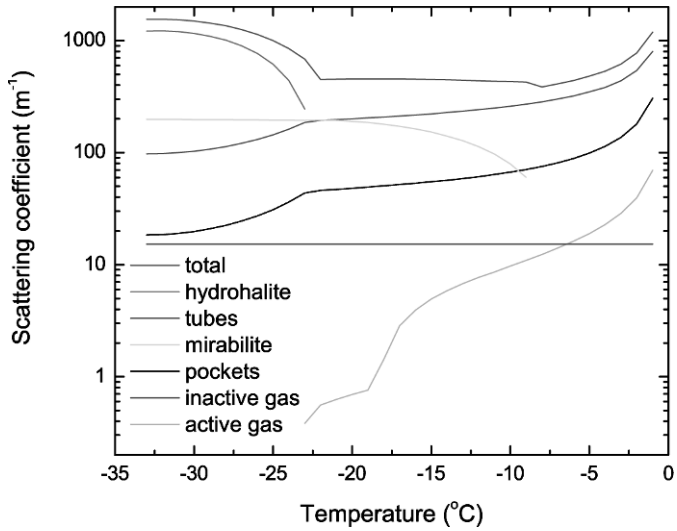


Fig. 8.21. Temperature-dependent scattering coefficients for various types of scatterers in sea ice as predicted by the structural-optical model. [For the color version of this figure, see the color plate section.]

8.7.1.4 Gas bubbles

We designate two types of gas bubbles in sea ice: (i) *active* bubbles contained in brine inclusions whose size varies with changes in the size of the host brine inclusion, and (ii) *inactive* bubbles trapped in the ice lattice whose size remains approximately constant with temperature. As before, scattering by gas bubbles in brine inclusions was first computed directly from observed size and number distributions at $-15°$C. Somewhat unexpectedly, only a fraction of the brine inclusions were found to contain bubbles at this temperature. Estimates of the total number density of active bubbles N_{bubble} were based on the percentage of brine inclusions containing bubbles and on the gas volume fraction of these inclusions. The fraction of brine pockets (f_{bp}) and tubes (f_{bt}) containing bubbles was observed to be 10% and 40%, respectively. Typical fractional gas volumes for these brine inclusions were estimated to be 10% for pockets ($v_{\mathrm{bub\text{-}bp}}$) and 3% for tubes ($v_{\mathrm{bub\text{-}bt}}$) at $-15°$C. As the temperature of the sample increases, some of the pure ice (density $\approx 0.917\,\mathrm{g\,cm^{-3}}$) around each brine inclusion melts to become brine with density greater than $1.0\,\mathrm{g\,cm^{-3}}$. It was assumed that any existing gas bubble would expand to fill the difference in volume. Changes in gas bubble size with temperature thus depend on changes in the volume of the associated brine inclusion, information which can be obtained from the freezing-equilibrium relations (Cox and Weeks, 1983).

It should be noted, however, that temperature-dependent changes in the size of active bubbles also depend on the actual volume of the host brine inclusion. This is simply illustrated by considering two identical gas bubbles, one in a large brine inclusion and the other in a small brine inclusion. Upon warming, both brine inclusions increase their volume by the same percentage, meaning that a larger

volume of ice would melt around the large brine inclusion than the small one. Hence, the size of the bubble in the larger inclusion would increase more than the one in the smaller inclusion, tying the temperature-dependent behavior of the bubble size distribution to the brine inclusion size distribution. This effect is taken into account in this model. By specifying f_{bp}, f_{bt}, $v_{\text{bub-bp}}$, and $v_{\text{bub-bt}}$ at $-15°\text{C}$, the active bubble distribution can be manipulated to simulate enlarging and shrinking as the ice is warmed and cooled. Changes in the volume of an individual active bubble are

$$V_{\text{bubble-a}}(T) = V_{\text{bubble-a}}(-15°\text{C}) + \Delta V_{\text{brine}}\left(\frac{\rho_{\text{brine}}(T)}{\rho_{\text{ice}}} - 1\right), \tag{8.26}$$

where $V_{\text{bubble-a}}$ is the volume of the active bubble, ΔV_{brine} is the fractional change in volume of the host brine inclusion between $-15°\text{C}$ and T, as specified by freezing equilibrium; $\rho_{\text{brine}}(T)$ varies between $1.03\,\text{g}\,\text{cm}^{-3}$ at $-2°\text{C}$ and $1.23\,\text{g}\,\text{cm}^{-3}$ at $-34°\text{C}$, as given by Maykut and Light (1995); and the density of pure ice $\rho_{\text{ice}}(\text{g}\,\text{cm}^{-3}) = 0.917 - 1.403 \times 10^{-4}T$ (°C) (Pounder, 1965). In this treatment, the number density of active bubbles is assumed to remain constant with temperature.

For the purpose of calculating ψ and σ, bubbles are assumed to be spherical at $-15°\text{C}$. At higher temperatures, however, bubbles in tubes frequently become large enough that their spherical dimensions exceed the tube diameter. When the bubble diameter grows larger than the tube diameter, the diameter of the bubble becomes fixed at the tube diameter, the shape of the bubble is assumed to be cylindrical, and the bubble is permitted to grow in length only. For such tubular bubbles, $\psi_{\text{bubble-a}}$ and $\sigma_{\text{bubble-a}}$ are calculated using the equivalent spheres treatment, in the same manner as brine inclusions. The predicted $\sigma_{\text{bubble-a}}(T)$ for all active gas bubbles is shown in Fig. 8.21. Although much smaller in magnitude, it is more sensitive to temperature than either tubes or pockets, showing large increases in scattering as the temperature increases above $-10°\text{C}$.

With $f_{\text{bp}} = 0.1$, $f_{\text{bt}} = 0.4$, $v_{\text{bub-bp}} = 0.1$, and $v_{\text{bub-bt}} = 0.03$, the fractional volume of active gas in the optical sample is computed to be 2.4×10^{-4} at $-15°\text{C}$, about 2% of the total gas volume indicated by the density and salinity measurements. It appears that there must have been additional gas or void space in the optical sample that was not recorded in the structural observations. These inclusions are important because of their large refractive index contrast with the surrounding ice. Reexamination of the imagery used to characterize the microstructure revealed previously unidentified features that are probably inactive gas bubbles embedded in the ice, e.g. see features highlighted in the lower right-hand corner of Fig. 8.5. The effective diameters of these bubbles range from 0.09 to 0.21 mm with a number density of about $0.24\,\text{mm}^{-3}$, indicating that the value of ψ for inactive gas ($\psi_{\text{bubble-i}}$) at $-15°\text{C}$ is $7.6\,\text{m}^{-1}$ and σ for inactive gas ($\sigma_{\text{bubble-i}}$) at $-15°\text{C}$ is $15.2\,\text{m}^{-1}$ in this thin section. These values should remain constant with temperature, as assumed in Fig. 8.21.

The fractional gas volume from the inactive bubbles observed in this image was estimated to be 0.00215. This, combined with the volume of active gas, still only accounts for 25% of the total predicted gas volume. We will attempt to quantify the gas content of the optical sample more precisely when we present the comparison of optical and structural data in subsection 8.7.3.

8.7.1.5 Precipitated salt crystals

The structural-optical model considers both mirabilite crystals that form below $-8°C$ and hydrohalite crystals that begin to form at $-23°C$. Observed crystal sizes in the larger inclusions ranged from $15\,\mu m$ to $140\,\mu m$ diameter. In subsection 8.4.3.4, we estimated that $\psi_{\text{mirabilite}}(-15°C) = 48.7\,m^{-1}$ in the 3‰ thin section sample. This estimate was based on the simple assumption of an effective crystal edge length of $10\,\mu m$ since direct measurement of the size distribution was not attempted.

With a salinity of 4.7‰, the optical sample should contain $9 \cdot 10^{-4}\,kg$ of mirabilite per kg of sea ice at $-15°C$ (Richardson, 1976). Assuming a crystal edge length of $10\,\mu m$, $\psi_{\text{mirabilite}}(-15°C) = 76.3\,m^{-1}$ and $\sigma_{\text{mirabilite}}(-15°C) = 152.6\,m^{-1}$. The value of $\sigma_{\text{mirabilite}}(T)$ is shown in Fig. 8.21, where the temperature-dependent total mass was based strictly on freezing equilibrium, and the crystal edge length was held constant at $10\,\mu m$. The assumption of constant crystal size was made despite the observation that existing mirabilite crystals grew in size while additional crystals nucleated. This observation, however, does not preclude the possibility of a temperature-independent *effective* crystal size.

Fig. 8.21 suggests that changes in scattering due to mirabilite are roughly balanced by opposite changes related to brine inclusions and active gas bubbles. As the ice warms, mirabilite crystals dissolve and their scattering diminishes, while brine inclusions and active gas bubbles enlarge and their scattering is enhanced. Our experimental results suggest that this balance may be independent of ice type and details of the microstructure. If ice salinity is increased, for example, σ for all the inclusions present will increase due to greater brine volume and mass of precipitates, but relative changes with temperature still act to offset one another.

Although hydrohalite crystals should be present in brine inclusions when $T < -23°C$, they were apparently also too small to be seen in the thin section imagery. However, the increased scattering below $-23°C$ (Fig. 8.21) indicates that significant amounts of hydrohalite did precipitate in the optical sample. To calculate $\sigma_{\text{hydrohalite}}(T)$, we again assumed that temperature dependence was based strictly on the freezing equilibrium mass fraction and that crystal size remained constant at $10\,\mu m$ (Fig. 8.21). Because brine and gas volumes are small at the low temperatures where hydrohalite precipitates, the amount of scattering produced by brine tubes, pockets, and active gas bubbles is small relative to the scattering by hydrohalite crystals.

8.7.2 Phase functions

For inclusion sizes relevant to this study, g is approximately independent of inclusion size, but depends strongly on the real refractive index (m) of the inclusion relative to its environment. Relative refractive indices of brine in ice, gas in brine, mirabilite in brine, and hydrohalite in brine are shown in Fig. 8.22(a). Because m_{brine} depends strongly on temperature (Maykut and Light, 1995), g for an inclusion depends on temperature when brine is either the scatterer (e.g., brine in ice) or the background medium (e.g., gas or salt crystals in brine). For a particular inclusion size, the value of g is smallest when $|m - 1|$ is largest. In this case,

backscattering is largest for gas in brine, although gas bubbles in ice scatter almost as effectively. For hydrohalite and mirabilite crystals in brine pockets, $|m-1|$ is largest at their temperatures of initial precipitation since $m_{\text{mirabilite}}$ and $m_{\text{hydrohalite}}$ are assumed to be independent of temperature and m_{brine} increases with decreasing temperature. In fact, mirabilite crystals in brine have $|m-1| \approx 0$ at $-30°$C, so that they become essentially invisible at this temperature. Conversely, brine inclusions in ice reach their largest $|m-1|$ value at low temperature. As sea ice warms and the brine becomes less concentrated, $|m_{\text{brine}}-1|$ decreases, causing brine inclusions to have increasing forward scattering.

Mie theory was used to calculate phase functions for brine inclusions in ice, gas inclusions in ice, gas inclusions in brine, and precipitated salt crystals in brine, assuming that all could be represented by equivalent spheres. Representation of the phase function by its asymmetry parameter g (the cosine-weighted angular integral) permitted qualitative comparison of relative amounts of forward and backward scattering between different types of scatterers (Fig. 8.22(b)) and was used in the model to specify the Henyey-Greenstein scattering phase function. Typical inclusion sizes observed in the thin section were chosen for the calculation, although none of the predicted values were sensitive to size. Scattering by brine and salt inclusions is clearly very forward-peaked (larger g), while gas bubbles produce considerably more backscattering (smaller g). Based on these calculations, desalinated, bubbly ice would be expected to have a bulk g value in the vicinity of 0.86, whereas the first-year interior ice used in this study had a representative g value in the 0.98–0.99 range. Fig. 8.22(c) shows $(1-g)$, which is proportional to s. It can be seen that brine inclusions in ice contribute more than 5 times as much to s at $-33°$C as at $-1°$C, and that $(1-g)$ attributable to brine inclusions decreases by a factor of two between -8 and $-2°$C. Mirabilite in brine shows an order of magnitude decrease in $(1-g)$ between -8 and $-23°$C, while total changes for hydrohalite and gas in brine are less than a factor of two.

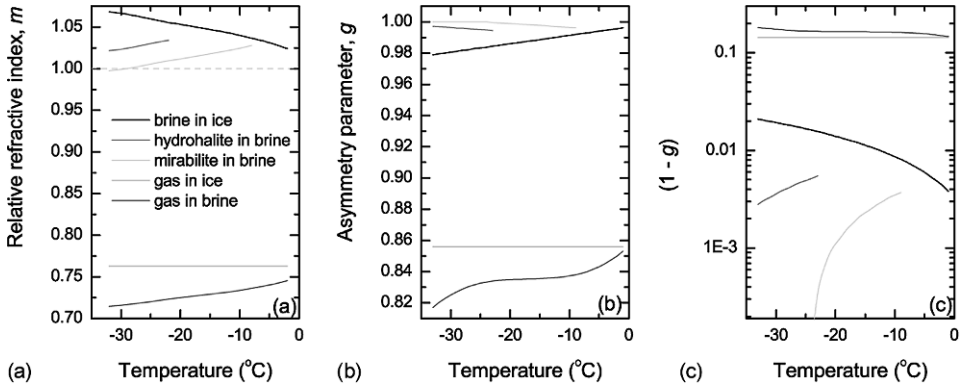


Fig. 8.22. Temperature-dependence of (a) relative refractive index, (b) asymmetry parameter, g, and (c) $(1-g)$ for inclusions of brine in ice, gas in ice, gas in brine, mirabilite in brine, and hydrohalite in brine. Asymmetry parameters were calculated using equivalent spheres and Mie theory. [For the color version of this figure, see the color plate section.]

8.7.3 Model development and testing

The relationships between ice physical properties and IOPs presented above form the basis for the structural-optical model. We will now discuss the formulation of this model, evaluate it for a preliminary set of conditions, and describe how it was tested and refined.

8.7.3.1 Preliminary model

The model is based on Eq. (8.23), which describes the spectral absorption within sea ice, and Eq. (8.25), which describes the contribution to scattering for each type of inclusion present. The total scattering coefficient is determined by

$$\sigma_{\text{total}} = \sum \sigma_i = \sigma_{\text{bp}} + \sigma_{\text{bt}} + \sigma_{\text{bubble-a}} + \sigma_{\text{bubble-i}} + \sigma_{\text{mirabilite}} + \sigma_{hydrohalite}. \quad (8.27)$$

Single-term Henyey–Greenstein phase functions were used to represent scattering phase functions for all scatterers in the ice. Other, possibly more realistic phase functions exist (e.g., Mobley et al., 2002), but the comparisons in this study have been restricted to the Henyey-Greenstein function. The temperature-dependent ensemble asymmetry parameter can be predicted by

$$g_{\text{total}}(T) = \frac{\sum_i \sigma_i(T) g_i(T)}{\sum_i \sigma_i(T)}, \quad (8.28)$$

where $g_i(T)$ are the asymmetry parameters for each type of scatterer as shown in Fig. 8.22(b). Equations (8.27) and (8.28), along with Eqs (8.23) and (8.25) constitute the structural-optical model.

8.7.3.2 Initial test

The model was tested by comparing values of $s(T)$ predicted by the model with those derived from the optical data (Fig. 8.20). Modeled $s(T)$ values were calculated from

$$s(T) = \sigma_{\text{total}}(T) \cdot (1 - g_{\text{total}}(T)) \quad (8.29)$$

using Eqs (8.23), (8.25), (8.27), and (8.28).

Fig. 8.23 shows the first attempt to compare predicted values of $s(T)$ with those derived from the optical observations. Significant discrepancies between the two exist in all three temperature regimes. An initial assessment suggests that the discrepancies occur for different reasons in the three regimes. For $T < -23°\text{C}$, this preliminary treatment significantly underestimates scattering, most likely by hydrohalite crystals. In the intermediate regime where no hydrohalite exists, it appears that improvements in the treatment of scattering by brine inclusions, gas bubbles, and mirabilite crystals are needed. This is not surprising, given that the model accounts for only 25% of the total gas volume in the sample and gas bubbles are highly effective scatterers. It also appears that some important physics may have been omitted when $T > -8°\text{C}$ since the model predicts large increases in scattering due to rapid enlargement of brine and gas inclusions while the optical observations show surprisingly little change.

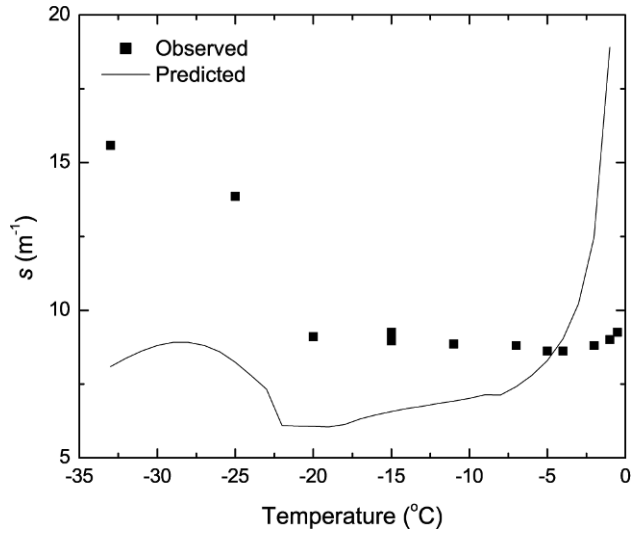


Fig. 8.23. Temperature-dependence of observationally determined similarity parameter compared with values predicted using the preliminary model.

8.7.3.3 Model refinements

This section seeks to improve the preliminary model by addressing problems related to uncertainties in gas volume, scattering by hydrohalite, and inclusion merging.

Gas volume

The most immediate problem is the estimate of the gas volume of the optical sample, as this potentially impacts the prediction of s at all temperatures. Imagery of the microstructure clearly shows bubbles within brine inclusions and bubbles embedded directly in the ice (Fig. 8.5). The arrows point out what appear to be drained brine tubes. At the time of initial analysis, it was assumed that the brine from these tubes had drained during thin section preparation. It is likely, however, that some of these tubes drained when horizontal cuts were made to remove both the thin section and the optical section from the ice core. Drained tubes present in the optical sample should behave as elongated inactive gas bubbles, with negligible temperature-dependent changes in size or scattering. Although these drained tubes might not be normally found in natural ice, they must be taken into account in predicting the optical properties of this sample.

Scattering by drained tubes was modeled by simply adding gas-filled tubes to the ice. The equivalent spheres treatment was applied to the drained tubes in exactly the same way as to the brine-filled tubes. Excellent agreement between the predicted and observed s was obtained at $-15°$C when the number of drained tubes was set to 3% of the number of brine tubes (Fig. 8.24). This contribution accounted for additional gas volume, bringing the total estimate of gas volume at $-15°$C to 30% of the volume calculated directly from the density, salinity, and temperature. If the bulk density of the sample were 0.921 g cm^{-3}, instead of the

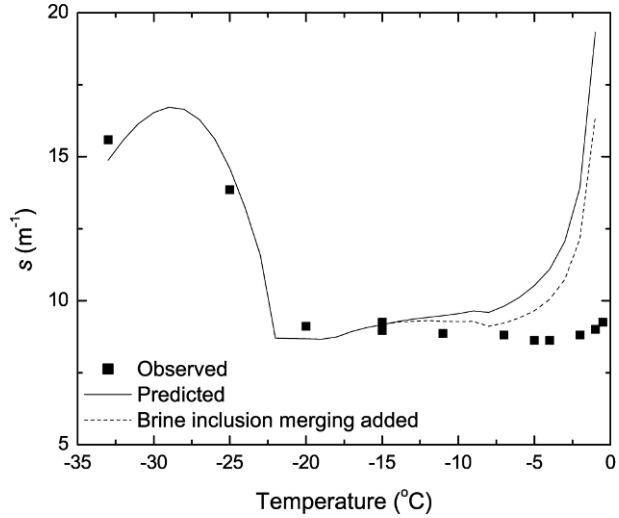


Fig. 8.24. Comparison of observationally determined and predicted values of $s(T)$. A revised model (solid line) includes parameterization of drained tubes and a reduction in the effective size of hydrohalite crystals from 10 to 4.5 μm. The dashed line shows the result of adding the effects of brine inclusion merging to these improvements.

observed $0.915 \pm 0.02\,\mathrm{g\,cm^{-3}}$, then the gas bubble parameterization would explain 98% of the predicted gas volume. This density is within the limits of measurement uncertainty.

For comparison, Mobley et al. (1998) estimated the amount of scattering within first-year ice sampled *in situ* off Pt. Barrow, Alaska, at the onset of melt. Their estimates for interior ice between the depths of 0.1 and 1.61 m produced an s value of $4.0\,\mathrm{m^{-1}}$. This value is considerably smaller than our observed values, and the difference may be attributable to the considerable brine drainage that occurred in our laboratory sample. Their measurements were made where the interior ice temperature was $-5.7°\mathrm{C}$, but the ice had salinity of 5.2‰ and density $0.92\,\mathrm{g\,cm^{-3}}$. These physical property measurements are consistent with a general lack of brine drainage.

Scattering by hydrohalite

Even with increased gas volume, the model still fails to produce enough scattering when $T < -23°\mathrm{C}$. To match the observations, it is necessary to have increased scattering by the hydrohalite crystals. Details of hydrohalite precipitation patterns within sea ice have not been investigated. So, as temperature decreases, the model assumes increasing numbers of crystals with a constant effective size of 10 μm. Newly nucleated crystals are presumably very small so it is not unreasonable that the effective size may be considerably less than assumed. By decreasing effective size from 10 μm to 4.5 μm, the associated increase in the surface area of hydrohalite causes $\psi(-33°\mathrm{C})$ to increase from $609\,\mathrm{m^{-1}}$ to $1354\,\mathrm{m^{-1}}$ and provides a reasonable match to the optical observations (Fig. 8.24). Keeping effective size constant is only one of several possible ways to treat this problem, but more information from

higher resolution observations of hydrohalite precipitation in sea ice is needed to develop better treatments.

Brine inclusion merging

Even with the improvements described above, there remains a significant discrepancy between predictions and observations at temperatures above −5°C. One likely reason is changes in the distribution of brine and gas inclusions that occurred in the samples as they approached the melting point. Imagery made at high temperatures showed that brine inclusions often merged with neighboring inclusions as they became larger. In some cases, large brine tubes merged with small neighboring inclusions. In other cases, clusters of small pockets merged to form a single inclusion. The net effect of this merging is a reduction in the total surface area of brine–ice interfaces and a corresponding reduction in the ψ values for brine pockets and tubes.

The effects of merging were included in the model by specifying a reduction in $N_{\text{brine}}(l)$. This was done by: (i) calculating the normal increase in the size of each brine inclusion with temperature, (ii) reducing the number density for each size to simulate merging, and (iii) increasing $l_{\max}$, the upper limit of the size distribution, to conserve total brine volume after merging. Specifically, $N_{\text{brine}}(l)$ at −15°C was first converted to a volume (v) distribution $N_{\text{brine}}(v, -15°\text{C})$ using the assumptions about shape and aspect ratio described in subsection 8.4.3.1. The temperature was then adjusted to T' and a new volume distribution $N_{\text{brine}}(v, T')$ calculated using the freezing equilibrium relations of Cox and Weeks (1983). Merging was simulated by

$$\tilde{N}_{\text{brine}}(v, T') = \eta(v, T') N_{\text{brine}}(v, T'), \tag{8.30}$$

where $\tilde{N}_{\text{brine}}$ is the merged distribution and η is a specified function based on observed merging data. Since merging was never observed below −14°C or in the smallest class of pockets, it was assumed that $\eta(v, T \le -14°\text{C}) = \eta(v_{\min}, T) = 1$, where $v_{\min}$ is the smallest volume in the distribution. The form selected for $\eta(v, T)$ was

$$\log[\eta(v, T)] = \frac{\log(\eta[v_{\max}, T])}{\log(v_{\max}/v_{\min})} \log\left(\frac{v}{v_{\min}}\right), \quad \text{for } v_{\min} < v < v_{\max} \tag{8.31}$$

where $v_{\max}$ is the volume of the largest inclusion at temperature T. $\eta(v, T)$ is linear in $\log(v) - \log(N)$ space (see Fig. 8.25), and causes most of the merging to occur in the larger inclusions. $\eta(v_{\max}, T)$, itself, decreases linearly with T, going from 1.0 at −14°C to 0.1 at −1°C, specifically, $\eta(v_{\max}, T) = 0.0307 - 0.0693T$, where T is in degree Celsius. While this particular parameterization reproduces the general merging behavior observed in the thin section, it is likely that better treatments can be developed as more data become available.

To conserve total brine volume in the sample, $\tilde{N}_{\text{brine}}(v, T')$ was extrapolated to larger volumes using the slope of the curve at v_{max}. An example of the effect of merging and volume conservation at $T = -1°\text{C}$ is shown in Fig. 8.25. The kink in the curves between $v = 10^{-3}$ and $10^{-2}\,\text{mm}^3$ marks the transition from pockets to tubes. The predicted effects of merging on $s(T)$ are shown by the dashed line in Fig. 8.25. Clearly, the merging model reduced scattering somewhat at the highest

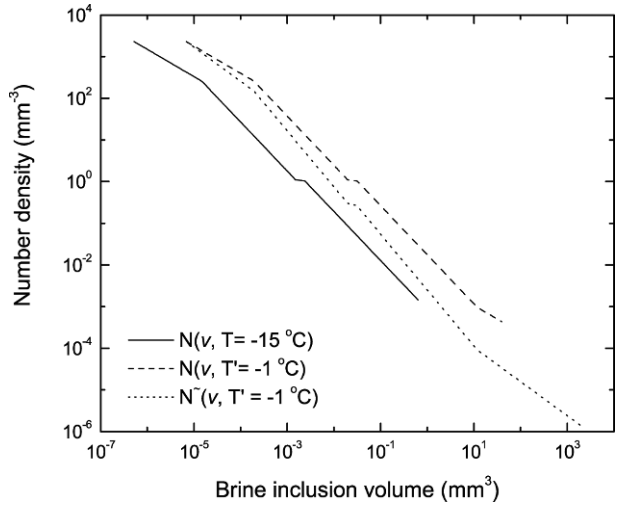


Fig. 8.25. Brine inclusion number distribution as a function of inclusion volume. $N_{\mathrm{brine}}(v, -15°\mathrm{C})$ was calculated from $N(l)$ (solid line), while $N_{\mathrm{brine}}(v, -1°\mathrm{C})$ (dashed line) was derived from $N_{\mathrm{brine}}(v, -15°\mathrm{C})$ using the freezing equilibrium equations. The merged distribution $\tilde{N}_{\mathrm{brine}}(v, -1°\mathrm{C})$ (dotted line) conserves total brine volume and assumes that $\eta(v_{\max}, -1°\mathrm{C}) = 0.1$. The area under the dashed and dotted curves is identical, as the merging process conserves total brine volume.

temperatures, but some other mechanism is still needed to explain why scattering does not increase strongly above $-5°\mathrm{C}$.

Gas bubble merging and escape

The rapid increase in brine volume above $-5°\mathrm{C}$ is also accompanied by proportional increases in the volume of active gas bubbles. Unfortunately, observational data are not sufficient to accurately characterize temperature-dependent changes in the distribution of active bubbles. The most probable explanation for the low scattering above $-5°\mathrm{C}$ is the upward escape of bubbles as the brine inclusions became larger and increasingly interconnected. Bubbles rising in melt ponds and vapor accumulation beneath the surface of refrozen ponds provide evidence that outgassing is a normal occurrence in warm sea ice, but the extent to which bubble escape from *in situ* ice is analogous to what happened in the laboratory sample is uncertain. Lacking specific information about the escape process, we simply assumed that all gas bubbles above a certain critical size (r_c) were able to escape from the sample. A very good fit to the optical data was obtained with $r_c = 0.16\,\mathrm{mm}$. Fig. 8.26 shows a comparison between the observed $s(T)$ and values predicted by the full model. Modifications made to the preliminary model produce a full model that closely matches the observed $s(T)$. This full model contains the essential elements of the physical changes observed in the microstructural observations, including temperature-dependent brine inclusion and gas bubble size, changes in brine concentration prescribed by freezing equilibrium, precipitation of solid salt crystals, brine tube drainage, brine inclusion merging, and escape of active bubbles with

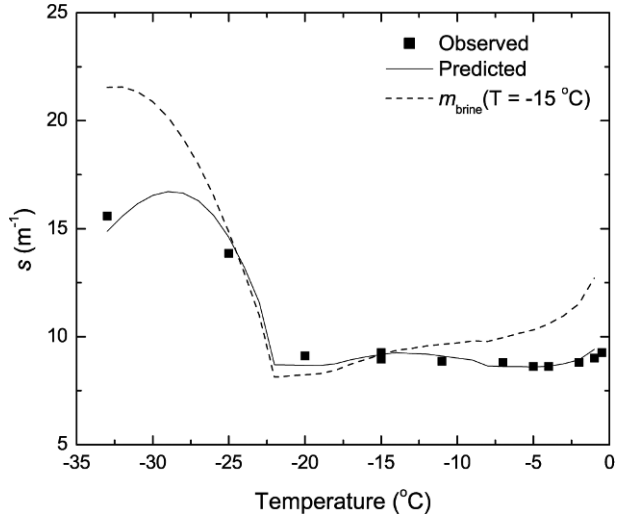


Fig. 8.26. Comparison of observationally determined $s(T)$ values (points) with those predicted by the full model with m_{brine} a function of temperature (solid line) and m_{brine} held fixed at its value at $-15°$C (dashed line).

critical radius $r_c = 0.16$ mm. The relative roles of σ and $(1 - g)$ will be discussed in subsection 8.7.4, along with sensitivities that affect the model.

8.7.4 Discussion

Relationships between the structural and optical properties of sea ice vary with ice temperature and are governed by different constituent scatterers at different temperatures. Fig. 8.27(a) shows the contributions due to partial $s(T)$ values for each type of scatterer based on the refined model. The values of $s(T)$ in this figure corroborate the optical observations for each temperature regime. Partial s values indicate that scattering at $T < -23°$C is dominated by hydrohalite crystals (magenta curve). In contrast, moderate temperatures are characterized by scattering produced by brine (red and black curves), gas (green curve), and mirabilite crystals (cyan curve). No one constituent dominates changes in scattering at these temperatures. Increases in the size of brine inclusions and active gas bubbles cause scattering coefficients to increase as the ice warms. Dissolution of mirabilite crystals upon warming causes the scattering coefficient for the crystals to decrease. Scattering in this regime remains nearly independent of temperature, the result of an approximate balance between enlarging brine and gas inclusions and dissolving mirabilite crystals. For $T > -8°$C, s for active gas bubbles increases rapidly, and this is manifested in slightly increased scattering in the observed optical properties.

Partial s values for each constituent depend on both the ice microstructure (σ) and its dielectric properties $(1 - g)$. Both the microstructure and the dielectric properties are inextricably tied to the chemistry of freezing equilibrium in seawater. Freezing equilibrium dictates the bulk amounts of ice, brine, gas, and precipitated salt that exist at each temperature in the seawater system, but does not dictate a

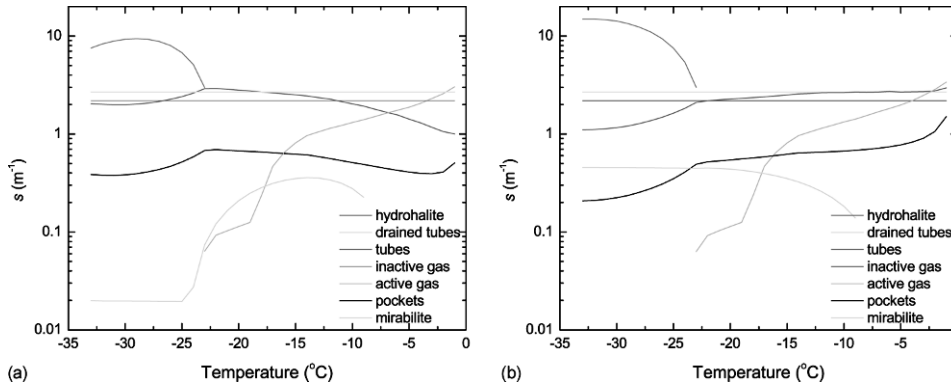


Fig. 8.27. Temperature-dependent partial similarity parameters for each constituent in the full structural-optical model using (a) $m_{\text{brine}}(T)$ and (b) $m_{\text{brine}}(T = -15°\text{C})$. [For the color version of this figure, see the color plate section.]

specific distribution of inclusions. A model that depends strictly on the chemistry of freezing equilibrium could be the most widely applicable type of structural-optical model. If such a model were appropriate, the optical properties of the ice could be determined without explicit knowledge of ice type, age, condition, or specific microstructure. To assess the applicability of the structural-optical model developed in this study, we consider the effect of processes governed strictly by freezing equilibrium relative to processes influenced by the ice type or its history. We begin by assessing the impact of $m_{\text{brine}}(T)$ on the full model. We then consider effects of variable ice salinity, and finally the role of specific microstructures.

8.7.4.1 Role of $m_{\text{brine}}(T)$

Changes in s that result from variations of m_{brine} with temperature occur regardless of ice type or the exact distribution of inclusions. To illustrate the impact of the temperature-dependent dielectric properties of brine on the optical properties of sea ice, Fig. 8.27b indicates predicted s values where $m_{\text{brine}}(T)$ was assigned the constant value $m_{\text{brine}}(T = -15°\text{C})$.

Comparison of Figs 8.27(a) and 8.27(b) shows that taking into account the temperature dependence of m_{brine} significantly enhances s for brine pockets and tubes at temperatures below $-15°$C and significantly decreases their s values at temperatures higher than $-15°$C. Because the refractive index of mirabilite and hydrohalite is generally larger than $m_{\text{brine}}(T)$, the effect of scattering by these salts diminishes as the ice is cooled below $-15°$C. Because $|m - 1|$ for gas in brine is relatively large, temperature-dependent changes in m_{brine} have negligible impact on scattering by gas bubbles. The dashed line in Fig. 8.26 indicates that without the full representation of $m_{\text{brine}}(T)$, our model would significantly over-predict the temperature dependence of $s(T)$ for the laboratory sample. While it is likely that s for $T < -23°$C is more sensitive to the effective size of precipitated crystals than to the exact representation of $m_{\text{brine}}(T)$ at the highest temperatures, the effect of $m_{\text{brine}}(T)$ is to dramatically reduce the efficacy of scattering by brine

inclusions. This results in a gradual lessening of importance of the microstructure and a gradual increasing of importance of $m_{\text{brine}}(T)$ with increased temperature.

While a model formulated strictly on the temperature-dependent changes in $m_{\text{brine}}(T)$ does a reasonable job of predicting the temperature-dependence of structural-optical properties in sea ice, and would be attractive for its wide applicability to all types of sea ice, it specifically fails to account for the total magnitude of s. The role of ice salinity in determining the magnitude of s will be addressed next.

8.7.4.2 Role of ice salinity

Since the bulk ice salinity dictates brine volume and the total mass of precipitated salt, the overall magnitude of the scattering is expected to depend strongly on salinity. We first present application of the model to a high-salinity ice sample, and then discuss the sensitivity of the model to ice salinity.

Independent optical data that can be used for testing the structural-optical model were also collected in the laboratory. These data were taken from 10 cm thick samples of high-salinity, laboratory-grown sea ice. The ice was grown at $-20°\text{C}$ with average salinity 15‰ and a bulk density of $0.90\,\text{g}\,\text{cm}^{-3}$. Although detailed structural data were not collected, apparent optical properties were measured and the Monte Carlo radiative transfer model was used to derive $s(T)$ in the same manner as was done for Fig. 8.20. Fig. 8.28(a) shows the observations (points) for this high-salinity sample. Note that s values for the 15‰ sample are approximately 3 times larger than s values for the 4.7‰ sample.

The structural-optical model was run for the high-salinity sample. The brine inclusion size distribution was extended to include tubes as large as $l = 45\,\text{mm}$ to account for the additional brine volume. While larger brine tubes were added to the distribution, the number density of smaller brine inclusions was the same as used to model the 4.7‰ ice. Scattering by precipitated salt crystals was adjusted for the

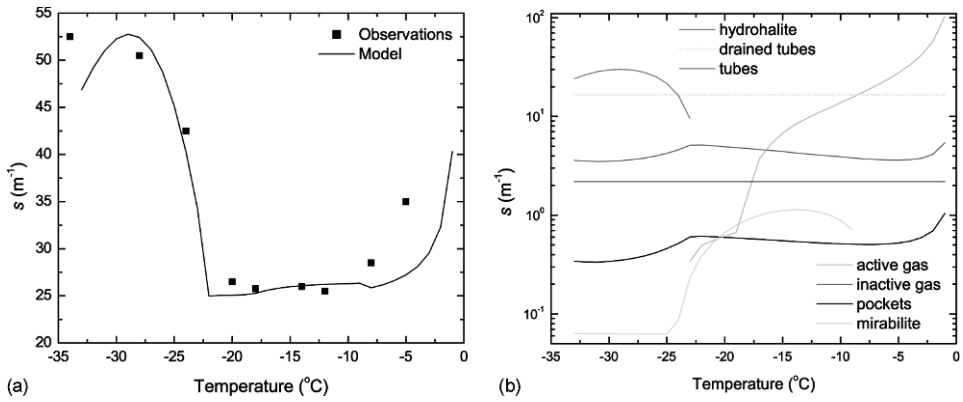


Fig. 8.28. Observations and model predictions for a high salinity sea ice sample, (a) comparison between s derived from observations and predicted with model, (b) model predicted similarity parameters for each constituent. [For the color version of this figure, see the color plate section.]

increased salinity by increasing the crystal number density. Effective crystal sizes inferred from the low-salinity case (10 μm for mirabilite and 4.5 μm for hydrohalite) were used, assuming that salinity does not affect crystal size, although this has not been specifically studied. The active gas bubble size and number parameterizations adopted for the low-salinity case were also used for this ice. As with the lower-salinity case, this parameterization produced only 30% of the total gas volume necessary to explain the observed bulk density. Also, as with the lower-salinity case, the bulk density would have had to been larger ($0.923\,\mathrm{g\,cm^{-3}}$) to be fully consistent with our bubble parameterization.

To achieve agreement with the optical observations, the percentage of drained tubes had to be increased from 3% to 14% for the high-salinity case. This is reasonable given that the brine volume of the high-salinity sample was three times larger than the lower-salinity case, likely leading to significantly larger brine drainage.

The solid line in Fig. 8.28a shows model predicted $s(T)$ for the high-salinity sample. The model explained observed scattering in the mirabilite regime and at $-28°$C in the hydrohalite regime. As in the lower-salinity case, hydrohalite crystals dominate scattering at low temperatures. Attenuation by scattering in this temperature regime depends on the size, number, and total mass of crystals, rather than their exact arrangement in the ice or the size and shape of the host brine inclusions. The optical properties of cold ice are thus controlled primarily by the precipitation patterns of the hydrohalite crystals, their dielectric properties, and the bulk ice salinity. However, scattering was under-predicted at $-34°$C, suggesting that the assumption of constant crystal size may need to be modified. The high-salinity ice had greatly increased scattering above $-8°$C which the model was unable to predict, even when gas bubble escape and brine inclusion merging were disabled. Since brine was observed to leak out of the sample at high temperatures, it is likely this increased scattering was caused by additional brine tube drainage.

In the mirabilite regime, the results suggest that the balance between increased scattering by brine and gas inclusions with warming and decreased scattering by mirabilite crystals still holds, despite the enhanced salinity. Again, it appears this result is fairly insensitive to the details of the microstructure, so that scattering in a wide variety of other sea ice types is also likely to exhibit weak temperature dependence in the mirabilite regime. The bulk salinity appears to affect the magnitude of total scattering, while the temperature-dependence is more dependent on $m_{\mathrm{brine}}(T)$ and, possibly, on the microstructure.

8.7.4.3 Role of microstructure

A number of model simulations were carried out to examine how changes in salt crystal, gas bubble, and brine inclusion distributions affect structural-optical relationships in sea ice.

Salt crystal size and number

Effective crystal sizes in the full model were assumed to be 10 μm for mirabilite, and 4.5 μm for hydrohalite, both independent of temperature. The crystal number is constrained by the total precipitable mass, as determined by freezing-equilibrium.

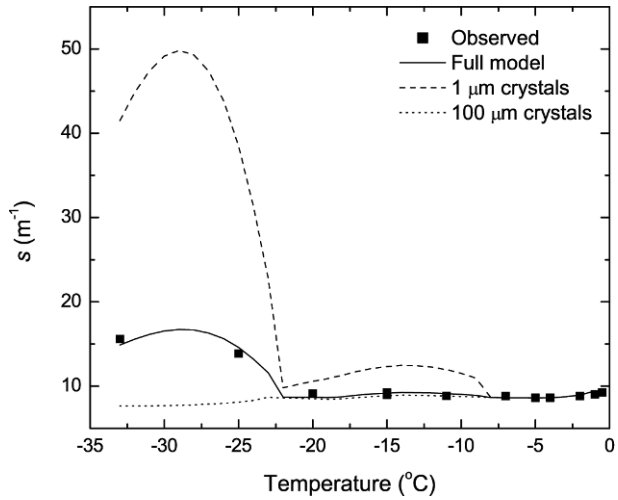


Fig. 8.29. Effects of precipitated crystal size on the temperature-dependence of the similarity parameter.

Fig. 8.29 shows the sensitivity of $s(T)$ to crystal size. Simulations were run with effective crystal size increased to 100 μm (dotted line) and decreased to 1 μm (dashed line). Increased crystal size had little impact on $s(T)$ for mirabilite since scattering by 10 μm crystals was already relatively small, whereas increasing the hydrohalite size from 4.5 μm to 10 μm dramatically decreased scattering at the lowest temperatures. Decreasing effective crystal sizes to 1 μm significantly enhanced σ for each type of crystal. However, despite significant increases in σ, effects of $m_{\text{brine}}(T)$ still reduced attenuation by scattering for mirabilite at temperatures below -15°C and for hydrohalite below about -29°C.

There may likely be other factors that affect scattering by precipitated salt crystals. Perhaps the most outstanding uncertainty is the prediction of single scattering phase functions for individual crystals. The calculations of phase functions for precipitated salt crystals were done using Mie theory, modeling the crystals as equivalent spheres, ignoring effects of the faceted crystal geometry. Both crystals form in the monoclinic system and have planar faces and sharp edges. We would expect the explicit treatment of single scattering by such crystals to produce considerably smaller values of g, commensurate with smaller inferred values of σ and larger inferred effective size, but we are not aware of single scattering calculations specifically for mirabilite and hydrohalite crystals in brine. There is also some evidence that ice crystals may co-crystallize with hydrohalite within the inclusions, further complicating the characterization of scattering at low temperatures.

Gas bubble size and number

The sensitivity to active gas bubbles was investigated by first doubling their fractional volume in brine inclusions, then by doubling the number of brine inclusions containing bubbles. Both tests doubled the total active gas volume. Doubling the volume of individual bubbles (Fig. 8.30, dashed line) produced only slightly enhanced s values in the mirabilite regime. There was also little effect in the warm

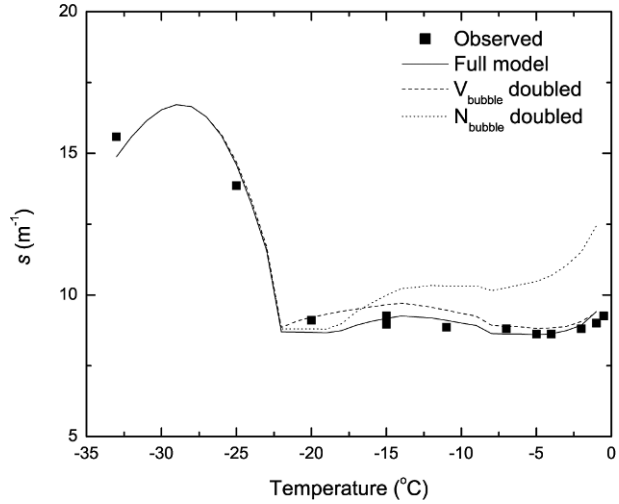


Fig. 8.30. Temperature-dependent changes in the similarity parameter as a function of active gas bubble volume and number density.

regime because bubbles with radius $r > 0.16\,\text{mm}$ were still assumed to escape from the ice. Doubling the number of active gas bubbles (Fig. 8.30, dotted line), however, strongly increased scattering. Doubling the number means that 80% of all brine tubes would contain bubbles at $-15°\text{C}$, contrary to the microstructural observations of 40%. The optical properties are more sensitive to the number simply because σ doubles when N doubles, but when the volume of individual inclusions is doubled, σ increases by only 1.59. At this point, the behavior and nucleation patterns of gas bubbles within brine inclusions remain poorly understood but are clearly important.

Distribution of brine

In this study, scatterers have been treated as being uniformly distributed and randomly oriented, a condition for the application of equivalent spheres. There is, however, ample evidence that the microstructure of first-year sea ice is generally oriented, such that brine tubes and strings of brine pockets are generally aligned or stacked vertically (e.g., Cole and Shapiro, 1998; Light et al., 2003a; Ehn et al., 2008). Long, straight-walled tubes with perfect vertical alignment should predominantly scatter light by azimuthal redistribution, which would not modify the irradiance attenuation. The structural alignment thus potentially impacts treatment of the IOPs for the microstructure.

Furthermore, it is likely that much of the *in situ* laboratory brine drainage may not ordinarily occur within pack ice. Eliminating the scattering contribution from drained tubes, would reduce $s(-15°\text{C})$ from 9.2 to $6.6\,\text{m}^{-1}$. Drained tubes do not respond to changes in temperature, and thus do not significantly impact the temperature response of $s(T)$. Gas bubble escape was also a key parameterization in developing the full model. Clearly, if gas bubbles fail to escape from the ice, particularly from the interior, the model shows that $s(T)$ would increase rapidly at high

temperatures. This could have important implications for the optical properties of melting sea ice, but not enough is currently known about gas bubbles in ice to properly quantify under what conditions they would escape or be retained.

8.8 Conclusions

The structural-optical model developed in this study is a useful tool for predicting inherent optical properties of first-year sea ice directly from information about ice temperature, bulk salinity, and the distribution of brine, gas, and salt inclusions. Three temperature regimes were identified, each exhibiting a distinct relationship between the ice microstructure and its inherent optical properties. In general, sea ice exhibits strong forward scattering, with bulk g values ranging between 0.86 for fresh bubbly ice and 0.98–0.99 for first-year interior ice. Table 8.2 gives a summary of suggested values of s, σ, and g for each of the three temperature regimes for the model developed in this work as applied to a sample of 4.7‰ first-year interior sea ice as observed under laboratory conditions (see Fig. 8.26).

Table 8.2. Representative range of s, σ, and g values for each of the three temperature regimes for the structural-optical model. Values are applicable to a sample of 4.7 ppt first-year, interior sea ice, observed under laboratory conditions.

Temperature regime	s (m^{-1})	σ (m^{-1})	g
Low: $T < -23°\mathrm{C}$	11.6–16.7	1002–3060	0.988–0.995
Medium: $-23°\mathrm{C} < T < -8°\mathrm{C}$	8.7–9.3	401–474	0.978–0.982
High: $T > -8°\mathrm{C}$	8.6–9.4	341–448	0.975–0.979

Since the temperature, salinity, and brine and gas distributions ultimately control the chemical, dielectric, and microstructural properties of sea ice, ties between the physical and optical properties of the ice are inextricably linked to freezing equilibrium. Freezing equilibrium impacts the optical properties of sea ice in two fundamental ways. Scattering coefficients for inclusions of brine, gas, and salt respond to changes in constituent volumes determined by freezing equilibrium. As brine inclusions and active gas bubbles grow and shrink and salt crystals precipitate and dissolve, scattering coefficients reflect changes in the total cross-sectional area of each population. Also, scattering phase functions for brine in ice, gas in brine, and salt crystals in brine respond to changes in the dielectric properties of brine. In addition to these important temperature-dependent processes, inactive gas bubbles contributed significant scattering to the laboratory sample. Little is known about the distribution of inactive gas in sea ice, but it is clearly an important component of the total scattering, at least in the ice used in this study.

The AOPs of sea ice are surprisingly independent of the strong temperature-dependent changes that occur in the microstructure of warm ice. This behavior is likely due to the strong scattering by inactive gas bubbles, but is also attributable to a balance between increasing scattering coefficient and increasing g as the brine becomes more dilute and its refractive index decreases. There is an analgous balance

between increasing brine inclusion size and decreasing mirabilite crystal number density which also acts to minimize the temperature-dependence of the AOPs at moderate temperatures ($-23°\mathrm{C} < T < -8°\mathrm{C}$). At lower temperatures, poorly understood details of the hydrohalite precipitation control the AOPs, but these conditions are of little practical importance. Most cases of geophysical interest occur at $T > -23°\mathrm{C}$ where, because of the various balances, it should be possible to develop relatively simple AOP parameterizations suitable for use in large-scale models.

It should be noted, however, that the results reported here are based on data collected in the laboratory from interior first-year ice. While the magnitude of *in situ* optical properties are likely to be different than the laboratory values, the temperature-dependence should be similar, except at very high temperatures where the ice may exhibit significant differences in brine drainage (i.e., in the volume of inactive gas) and the loss of active bubbles. Likewise, little is known about the detailed microstructure of multiyear ice and the degree to which the structural-optical model can predict AOPs in such ice is uncertain. A recent study (Light et al., 2008) found that s values in multiyear ice typically decrease by a factor of 50 from the surface layers of the ice to its interior. Interpretation of IOPs in sea ice, however, is complicated by strong forward scattering and varying degrees of structural anisotropy inherent to the ice. The extent of our quantitative understanding of the structural-optical properties and processes that occur in sea ice is likely to increase with additional field observations and further modeling of these processes.

Acknowledgments

The author is grateful for the supervision of her dissertation advisors Dr Gary Maykut and Dr Thomas Grenfell under whom this work was completed. Financial support from the United States Office of Naval Research, High Latitude Dynamics Program and the Office of Polar Programs at the United States National Science Foundation is gratefully acknowledged.

References

Abramowitz, M. and I. A. Stegun (Eds.) (1965), *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York.

Adams, C. A. and G. W. Kattawar (1978), Radiative transfer in spherical shell atmospheres 1. Rayleigh scattering, *Icarus*, **35**, 139–151.

Adams, L. H. and R. E. Gibson (1930), The melting curve of sodium chloride dihydrate: An experimental study of an incongruent melting at pressures up to twelve thousand atmospheres, *J. Am. Chem. Soc.*, **52**, 4252–4264.

Alcouffe, R. E., R. S. Baker, F. W. Brinkley, D. R. Marr, R. D. O'Dell and W. F. Walters (1997), DANTSYS: A diffusion accelerated neutral particle transport code system, Los Alamos Natl. Lab, Univ. of Calif., Los Alamos, NM.

Arcone, S. A., A. J. Gow and S. McGrew (1986), Microwave dielectric, structural, and salinity properties of simulated sea ice, *IEEE Trans. Geosci. Rem. Sens.*, **24**, 832–839.

Case, K. M. (1957), Transfer problems and the reciprocity principle, *Rev. Mod. Phys.*, **29**, 651–663.

Case, K. M. and P. F. Zweifel (1967), *Linear Transport Theory*, Addison Wesley, Reading, MA.

Cashwell, E. D. and C. J. Everett (1959), *A Practical Manual on the Monte Carlo Model for Random Walk Problems*, Pergamon Press, New York.

Cole, D. M. and L. H. Shapiro (1998), Observations of brine drainage networks and microstructure of first-year sea ice, *J. Geophys. Res.*, **103**, 21739–21750.

Collins, D. G., W. G. Blattner, M. B. Wells and H. G. Horak (1972), Backward Monte Carlo calculations of the polarization characteristics of the radiation emerging from spherical-shell atmospheres, *Appl. Opt.*, **11**, 2684–2696.

Cox, G. F. N. and W. F. Weeks (1983), Equations for determining the gas and brine volumes in sea-ice samples, *J. Glaciol.*, **29**, 306–316.

Crosbie, A. L. and R. L. Dougherty (1981), Two-dimensional radiative transfer in a cylindrical geometry with anisotropic scattering, *J. Quant. Spect. Rad. Trans.*, **25**, 551–569.

Ehn, J. K., C. J. Mundy and D. G. Barber (2008), Bio-optical and structural properties inferred from irradiance measurements within the bottommost layers in an Arctic landfast sea ice cover, *J. Geophys. Res.*, **113**, *C03S03* doi: 10.1029/2007JC004194.

Eicken, H. (1993), Automated image-analysis of ice thin-sections – instrumentation, methods and extraction of stereological and textural parameters, *J. Glac.*, **39**, 341–352.

Eicken, H., C. Bock, R. Wittig, H. Miller and H. O. Poertner (2000), Magnetic resonance imaging of sea-ice pore fluids: methods and thermal evolution of pore microstructure, *Cold Reg. Sci. Tech.*, **31**, 207–225.

Evans, K. F. (1993), A general solution for stochastic radiative transfer, *Geophys. Res. Letters*, **20**, 2075–2078.

Gabriel, P. M. and K. F. Evans (1996), Simple radiative transfer methods for calculating domain-averaged solar fluxes in inhomogeneous clouds, *J. Atm. Sci.*, **53**, 858–877.

Gavrilo, V. P. and B. Y. Gaitskhoki (1970), The statistics of air inclusions in ice, in *The Physics of Ice*, 125–128, translated from Russian by the Israel Program for Scientific Translations, Jerusalem.

Gitterman, K. E. (1937), Termicheskiy analiz morskoy vody [Temperature analysis of sea-water]. *Trudy Solyanoy Laboratorii, Vyp*, **15**(1), 5–23.

Gordon, H. R. (1985), Ship perturbation of irradiance measurements at sea. 1: Monte Carlo simulations, *Appl. Opt.*, **24**, 4172–4182.

Graaff, R., J. G. Aarnoudse, F. F. M. de-Mul and H. W. Jentink (1993), Similarity relations for anisotropy scattering in absorbing media, *Optical Engineering*, **32**, 244–252.

Grenfell, T. C. (1983), A theoretical model of the optical properties of sea ice in the visible and near infrared, *J. Geophys. Res.*, **88**, 9723–9735.

Grenfell, T. C. (1991), A Radiative transfer model for sea ice with vertical structure variations, *J. Geophys. Res.*, **96**, 16991–17001.

Grenfell, T. C. and D. K. Perovich (1981), Radiation absorption coefficients of polycrystalline ice from 400–1400 nm, *J. Geophys. Res.*, **86**, 7447–7450.

Grenfell, T. C. and S. G. Warren (1999), Representation of a nonspherical ice particle by a collection of independent spheres for scattering and absorption of radiation, *J. Geophys. Res.*, **104**, 31697–31709.

Henyey, L. G. and J. L. Greenstein (1941), Diffuse radiation in the galaxy, *Astrophys. J.*, **93**, 70–83.

Herman, B. M., A. Ben-David and K. J. Thome (1994), Numerical technique for solving the radiative transfer equation for a spherical shell atmosphere, *Appl. Opt.*, **33**, 1760–1770.

Light, B. (1995), A structural-optical model of cold sea ice, M.S. thesis, Univ. Wash., Seattle, 1 March.

Light, B., H. Eicken, G. A. Maykut and T. C. Grenfell (1998), The effect of included particulates on the spectral albedo of sea ice, *J. Geophys. Res.*, **103**, 27739–27752.

Light, B., G. A. Maykut and T. C. Grenfell (2003a), Effects of temperature on the microstructure of first-year Arctic sea ice, *J. Geophys. Res.*, **108**(C2) 3051, doi: 10.1029/2001JC000887.

Light, B., G. A. Maykut and T. C. Grenfell (2003b), A two-dimensional Monte Carlo model of radiative transfer in sea ice, *J. Geophys. Res.*, **108**(C7), 3219 doi: 10.1029/2002JC001513.

Light, B., G. A. Maykut and T. C. Grenfell (2004), A temperature-dependent, structural-optical model of first-year sea ice, *J. Geophys. Res.*, **109**(C6), C06013 doi: 10.1029/2003JC002164.

Light, B., T. C. Grenfell and D. K. Perovich (2008), Transmission and absorption of solar radiation by Arctic sea ice during the melt season, *J. Geophys. Res.*, **113**(C3), doi: 10.1029/2006JC003977.

Marion, G. M., R. E. Farren and A. J. Komrowski (1999), Alternative pathways for seawater freezing, *Cold Reg. Sci. Tech.*, **29**, 259–266.

Marshak, A., A. Davis, W. Wiscombe and R. Cahalan (1998), Radiative effects of sub-mean free path liquid water variability observed in stratiform clouds, *J. Geophys. Res.*, **103**, 19557–19567.

Maykut, G. A. and B. Light (1995), Refractive-index measurements in freezing sea-ice and sodium-chloride brines, *Applied Optics*, **34**, 950–961.

Mobley, C. D. (1994), *Light and Water: Radiative Transfer in Natural Waters*, Academic Press, San Diego, CA.

Mobley, C. D. and L. K. Sundman (2003), Effects of optically shallow bottoms on upwelling radiances: Inhomogeneous and sloping bottoms, *Limnol. Oceanogr.* **48**, 329–336.

Mobley, C. D., G. F. Cota, T. C. Grenfell, R. A. Maffione, W. S. Pegau and D. K. Perovich (1998), Modeling light propagation in sea ice, *IEEE Trans. Geosci. and Rem. Sens.*, **36**, 1743–1749.

Mobley, C. D., L. K. Sundman and E. Boss (2002), Phase function effects on oceanic light fields, *Appl. Opt.*, **41**, 1035–1050.

Nelson, K. H. and T. G. Thompson (1954), Deposition of salts from sea water by frigid concentration, *J. Mar. Res.*, **13**, 166–182.

Oikarinen, L. (2001), Polarization of light in UV-visible limb radiance measurements, *J. Geophys. Res.*, **106**, 1533–1544.

Perovich, D. K. (1990), Theoretical estimates of light reflection and transmission by spatially complex and temporally varying sea ice covers, *J. Geophys. Res.*, **95**, 9557–9567.

Perovich, D. K. and A. J. Gow (1991), A statistal description of the microstructure of young sea ice, *J. Geophys. Res.*, **96**, 16943–16953.

Perovich, D. K. and A. J. Gow (1996), A quantitative description of sea ice inclusions, *J. Geophys. Res.*, **101**, 18327–18243.

Perovich, D. K., C. S. Roesler and W. S. Pegau (1998), Variability in Arctic sea ice optical properties, *J. Geophys. Res.*, **103**, 1193–1208.

Porter, M. W. and R. C. Spiller (1956), *The Barker Index of Crystals; A Method for the Identification of Crystalling Substances*, W. Heffer, Cambridge.

Pounder, E. R. (1965), *The Physics of Ice*, Pergamon Press, Oxford.

Richardson, C. (1976), Phase relationships in sea ice as a function of temperature, *J. Glaciol.*, **17**, 507–519.

Ringer, W. E. (1928), Uber die Veranderungen in der Zusammensetzung des Meereswassersalzes beim Ausfrieren, *Rapports et Proces-verbaux des Reunions. Conseil Permanent International pour l'Exploration de la Mer*, **47**, 226–231.

Roberti, L. and C. Kummerow (1999), Monte Carlo calculations of polarized microwave radiation emerging from cloud structures, *J. Geophys. Res.*, **104**, 2093–2104.

Roberti, L., J. Haferman and C. Kummerow (1994), Microwave radiative transfer through horizontally inhomogeneous precipitating clouds, *J. Geophys. Res.*, **99**, 16707–16718.

Roedder, E. (1984), *Fluid Inclusions*, Mineral. Soc. of Am., Washington, DC.

Rozanov, A. V., V. V. Rozanov and J. P. Burrows (2000), Combined differential-integral approach for the radiation field computation in a spherical shell atmosphere: nonlimb geometry, *J. Geophys. Res.*, **105**, 22937–22942.

van de Hulst, H. C. (1980), *Multiple Light Scattering Tables, Formulas, and Applications*, Academic Press, New York.

Weeks, W. F. and S. F. Ackley (1982), The growth, structure and properties of sea ice, *CRREL Monograph 82-1*, Cold Reg. Res. and Eng. Lab., Hanover, NH.

Wu, S. C. and C. Y. Wu (1997), Radiative heat transfer in a two-dimensional cylindrical medium exposed to collimated radiation, *Int. Comm. Heat Mass Trans.*, **24**, 475–484.

Wyatt, C. L. (1978), *Radiometric Calibrations: Theory and Methods*, Academic Press, New York.

Zuidema, P. and K. F. Evans (1998), On the validity of the independent pixel approximation for boundary layer clouds observed during ASTEX, *J. Geophys. Res.*, **103**, 6059–6074.

9 Reflectance of various snow types: measurements, modeling, and potential for snow melt monitoring

Jouni I. Peltoniemi, Juha Suomalainen, Teemu Hakala, Jyri Näränen,
Eetu Puttonen, Sanna Kaasalainen, Manuela Hirschmugl and Johanna Torppa

9.1 Introduction

Seasonal snow covers large parts of the northern hemisphere annually. It can change the albedo of the surfaces from dark to bright overnight (and back), causing significant climate feedback (Manninnen and Stenberg, 2008; Flanner and Zender, 2006; Pirazzini, 2008; Nolin and Frei, 2001; Roesch et al., 2001). It forms large energy reservoirs which can be exploited by hydro energy power plants, and is the source of big floods when melting. It can significantly impact traffic and construction safety. It changes living and environmental conditions radically, and has major recreational value.

For all these, and many more purposes left unmentioned here, it is vital to observe the snow cover regularly (Matikainen et al., 2002; Metsämäki et al., 2002; Solberg et al., 1997). In crucial snow melt days, even daily estimates would be very much desired. Remote sensing techniques provide the only alternative for systematic large-area data collection, satellites can be used for global data and aeroplanes for regional data, and in the near future strato-balloons can fill the gap between aeroplanes and satellites, and small unmanned flying observation stations between aeroplanes and surface measurements.

The extent of snow cover is often estimated from images (Matikainen et al., 2002; Metsämäki et al., 2002, 2005; Salminen et al., 2009; Painter et al., 2009). While open areas are straightforward to interpret, the challenge is snow in forests. The albedo can be measured accurately, if sufficient multi-directional observations are available, but usually even in the best cases only very few (1–7) directions are seen, and thus some modeling and interpretation is needed (Nolin and Frei, 2001; Stroeve et al., 2005, 2006; Knap and Oerlemans, 1996; Manninen et al., 2006; Moody et al., 2007; Painter et al., 2009; Manninnen and Stenberg, 2008).

More details of snow properties can be revealed using multicolor observations or spectrometry (Green and Dozier, 1995, 1996; Nolin and Dozier, 2000; Painter et al., 2003a; Eskelinen et al., 2003; Tedesco and Kokhanovsky, 2007; Dozier et al., 2009; Lyapustin et al., 2009). Grain size, or effective size, or some parameter related to the scale of the snow layer can be inverted from the spectral shape (band ratios). The ambiguity lies in the definition of grain in a complicated 3-D structure of air and ice. Spectrometry can further yield information on impurities, soot content,

and algae. Of especial interest is the observation of wetness. Liquid water and ice are optically very similar in the optical regime. In only a few narrow bands do their absorption properties differ slightly, which has made it possible to try to estimate the melting.

To analyze all these data, and design optimal instruments, it is important to understand the reflective properties of the targets – snow and environment – to a high degree of accuracy. Ideally, one needs a physical reflectance model that predicts the observed signal for a given sensor in a given position as a function of the measurable physical and chemical properties of the target, e.g., grain size, wetness, density, thickness. Similarly, when inverted against reasonable wide observation data, this model should yield accurate estimates for the parameters.

The most popular way to model reflection of light by snow is to use the Mie scattering model together with the 1-dimensional radiative transfer theory (Bohren and Barkstrom, 1974; Wiscombe and Warren, 1980; Warren, 982; Leroux et al., 1998; Mishchenko et al., 1999; Jin and Simpson, 1999; Perovich, 1990; Green et al., 2002; Kokhanovsky and Zege, 2004; Kokhanovsky et al., 2005). Although the Mie/radiative transfer models can predict many spectral and directional features well, e.g., grain size dependence, it is clear by the empirical and theoretical data that the shape of the grains plays significant role (e.g., Mishchenko et al. (2000)), and the spherical Mie approximation cannot handle this. Some extensions for more general shape have been developed by Grenfell and Warren (1999) and Grenfell et al. (2005) using multi-sphere representation. Tanikawa et al. (2006) have also tried cylindrical and ellipsoidal particles. Peltoniemi (2007) has developed a model based on the stochastic grain shape model and ray-optics, which can already predict many shape and wetness related effects, but fails in some details. Another ray-trace model, ignoring polarizational effects but with an interesting 3-D modeling, has been developed by Kaempfer et al. (2007) as well.

To develop and validate the models further, a large set of reliable empirical measurements are needed. For as wide-ranging an analysis as possible, the full bidirectional reflectance factor (BRF) should be measured in the full hemisphere of incident and refracted radiation, in large spectral regime, including polarization, of a comprehensive set of targets all of whose relevant physical and chemical properties are separately measured. In practice, of course, one must be satisfied with a limited subset of that.

Measuring snow is challenging work, not only because the wet and cold snow environments require extra protection for instruments and measurement staff, or because of logistic complications, but also because snow itself is a difficult substance to deal with. It metamorphoses (changes all its key properties) continuously, sometimes in minutes, between many forms. It can have a complicated vertical structure, often also horizontal variations, with a random or wavy surface topography.

Middleton and Mungall (1952) pioneered reflectance measurement techniques in the 1950s with a specially made field goniophotometer. They measured several snow types under various illumination geometries using white broadband light. Since then new initiatives have been taken, e.g., by Kuhn (1985); Steffen (1987). Dozier, Nolin et al. continued measurements soon (Dozier et al., 1989; Nolin et al., 1990; Nolin and Dozier, 1991; Nolin et al., 1994) with aims for improved satellite data interpretation. More recent measurements have been made at least by Perovich

(1994); Leroux et al. (1998); Warren et al. (1998); Kaasalainen (2002); Green et al. (2002); Painter and Dozier (2002); Bonnefoy (2001); Giardino and Brivio (2003); Painter and Dozier (2004); Li and Zhou (2004); Hudson et al. (2006); Bourgeois et al. (2006a) using an improved spectral range and resolution while employing various techniques. These measurements have already found and confirmed spectral wetness signals around 1000 nm and 1300 nm, clear grain size dependence, and the effect of surface roughness. Peltoniemi et al. (2005a) also observed that there is a potential directional wetness signature near the specular direction, which had not yet been predicted by any model nor confirmed by other observations, and Peltoniemi et al. (2009) found a polarization signature in the forward direction. Since *in situ* reflectance measurements are now becoming a routine part of many observation campaigns, there are probably many more unpublished measurements or measurements briefly mentioned as parts of larger work.

In this chapter we review the reflection properties of snow, based on our large database of field measurements, presenting many new compositions of results, and comparing them to a model. Many of the results, conclusions and descriptions have already been published (Peltoniemi et al., 2005a; Peltoniemi, 2007; Peltoniemi et al., 2009), but a lot of additional material is also shown.

9.2 Snow

Snow flakes, or more precisely single ice crystals, are born in the atmosphere and subsequently fall down to earth where they create a cover of snow, a snow layer that can be semi-permanent or permanent. These single ice crystals, also known as snow crystals, grow from water vapor at freezing temperatures in the atmosphere and can reach sizes from hundreds of micrometers to millimeters. The snow crystal formation mechanism, although an extensively studied subject, has left large unanswered questions, even at a qualitative level (Libbrecht, 2005). The form the snow crystals take is a function of two parameters: the supersaturation of water and the temperature. Different forms include, e.g., dendrites, plates, columns, and prisms. For example, dendrites (classic snowflakes) grow at highly supersaturated environments in temperature regions of about 0 to -3 and -10 to -22°C. The classic text by Nakaya (1954) is a good introduction to the phenomenology of snow crystals.

A new snow layer has a relatively low density (5–15% of water) due to the porous nature of single snow crystals and the way they pack to form a snow layer. Snow layers consisting of highly branched or dendritic snow crystals tend to have a lower density and are often referred to as dry snow. Columnar or plate-like crystals will have much less air within the crystal structure and will therefore form denser snow layers that also feels wetter. Once the snow is on the ground, it will settle by its own weight until its density is ∼30% of that of water. This initial density can be changed by melting and freezing processes until it has reached a maximum of ∼50% of that of water for old snow.

When the snow layer starts to melt, part of the water is evaporated, part of it flows down, but the wet snow layer can hold about 5% of liquid water with ice. Where the liquid water exactly is stored in the snow layer is not very well

known yet, although different theories exist, including: as a coating over ice grains, between grains, in mini-ponds, and hanging under the grains. Some observations from melting crystals submerged in hexane is presented by Knight (1979).

Snow crystals metamorphose all the time, first growing in size and into complicated shapes and then, when hitting the ground, initially separating into needles and then sintering together to form larger grains. New ice can be formed by condensation of water vapor from air on old snow flakes, and the snow layer can grow in mass even without any new snow, or lose mass by evaporation. Normally, the grains grow to larger, irregularly rounded grains, but in certain conditions specific crystalline forms can form (Dingman, 2002).

The vertical structure in a snow layer can be rich. Especially in spring there is often a hard top layer some 1 to 10 cm thick that can withstand the weight of a skier or walker or sometimes even a vehicle. Below that can be several different layers of loose snow, deep hoar, slush, water, ice, more hard layers, etc., all metamorphosed in their own way. Vertical temperature gradients can be significant. The thickness of all the vertical layers vary horizontally, and actually there is also a rich 3-dimensional structure, holes, dunes, tracks, signs of underlying topography, vegetation, and many other objects over and under the snow surface.

A classification of various snow types is given, e.g., by Colbeck et al. (1985); Gray and Male (1981). Later at least the following snow types are referred: sd = stellar dendrites (snow flake), nd = needles, dc = partly decomposed, sr = small rounded, lr = large rounded, mx = mixed forms, cl = clustered rounded wet grains, mf = rounded melt–freeze polycrystals.

9.3 BRF, definitions

The bidirectional reflectance factor (BRF for short, denoted as R here) is defined as the ratio of the reflected light intensity of a given target to an ideal Lambertian reflector with a spherical albedo of 1.0 under the same incident irradiation (Hapke, 1993; Liang, 2004; Nicodemus et al., 1977)

$$R(\mu, \mu_0, \phi, \phi_0) = \frac{\pi I(\mu, \phi)}{\mu_0 F_0(\mu_0, \phi_0)}, \tag{9.1}$$

where F_0 is the incident collimated flux ($I_0(\Omega) = F_0\delta(\Omega - \Omega_0)$) and I the reflected radiance. The angles are defined in Fig. 9.1: ι and ϕ_0 are the zenith and azimuth angles of incidence, ϵ and ϕ are the zenith and azimuth angles of emergence, α is the scattering phase angle ($\cos\alpha = \cos\iota\cos\epsilon + \sin\iota\sin\epsilon\cos(\phi - \phi_0)$), a complement of the scattering angle. One can further define the specular direction ($= \iota, -\phi_0$), and the angle from that direction γ.

Above it was assumed, that the scattering is polarization neutral, or that the incident light is unpolarized and the sensor is polarization insensitive. While this applies satisfactorily for many cases – sunlight is unpolarized, ground surfaces polarize usually weakly, and most cameras see no polarization – there are enough counter cases to warrant more detailed treatment, e.g., blue skylight is strongly polarized, especially forward scattering can be rather polarizing or polarization sensitive, and polarimetry itself is an interesting field of research.

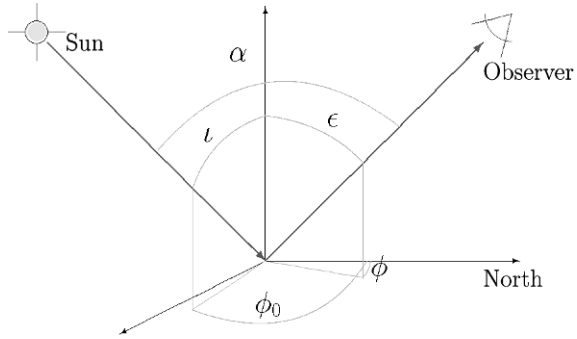


Fig. 9.1. Definition of the angles used in surface reflectance work: ϵ and ι are the zenith angles of the emergent (observer) and incident (solar) radiation respectively (shorthands $\mu = \cos\epsilon$ and $\mu_0 = \cos\iota$ are also used). ϕ and ϕ_0 are the corresponding azimuths. The phase or backscattering angle α is the angle between the observer and the sun. The principal plane is fixed by the solar direction and the surface normal, while the cross plane is a vertical plane perpendicular to the principal plane.

Polarization of electromagnetic radiation waves, such as light, is defined by the direction of the electric field. Natural light is a mixture of a very large number of various wave packets. Sunlight and many other light sources contain equal amounts of all polarization states, and are thus said to be unpolarized. Scattering, reflection, or transmission through asymmetric media often polarize light. The most significant polarization mechanisms are Rayleigh scattering (resulting in up to 90% polarization of blue sky) and Fresnel reflection (explaining the up to 100% polarization of water surfaces). However, all scattering mechanisms polarize to some extent, although the more complicated ones usually less so. One discussed subtlety is the small negative polarization near backward scattering angles caused by multiple scattering coherence effects (Muinonen, 2004; Mishchenko, 2008). The polarization properties depend mostly on the chemical composition (refractive index) and wavelength and sub-wavelength scale structures of the scatterers, and amount of multiple scattering.

A convenient way to describe natural light with arbitrary state of polarization is to use the Stokes parameters $[I, Q, U, V]$. hese are related to electromagnetic fields as

$$I \propto |E_\parallel|^2 + |E_\perp|^2, \tag{9.2}$$

$$Q \propto |E_\parallel|^2 - |E_\perp|^2, \tag{9.3}$$

$$U \propto E_\parallel {E_\perp}^* + {E_\parallel}^* E_\perp, \tag{9.4}$$

$$V \propto E_\parallel {E_\perp}^* - {E_\parallel}^* E_\perp, \tag{9.5}$$

where $E_\parallel$ and $E_\perp$ are electromagnetic field amplitudes parallel and perpendicular to some selected plane. It is important to note that the polarization plane must be defined in order to uniquely use the Stokes parameters. The definition of BRF is easily modified for polarization by replacing the scalar intensity by a Stokes vector $I \leftarrow \mathbf{I} = [I, Q, U, V]$, and $R \leftarrow \mathbf{R} = 4 \times 4$ Muller matrix.

One can further define the degrees of linear polarization as

$$P_Q = -\frac{Q}{I}, \tag{9.6}$$

$$P_U = -\frac{U}{I}, \tag{9.7}$$

$$P_P = \frac{\sqrt{Q^2 + U^2}}{I}, \tag{9.8}$$

which have the advantage that they are independent of the calibration of the sensor, and can thus be measured in conditions where reflectances cannot be normalized. When extracting degrees of polarization from spectral data for wider band simulation one must integrate nominator and denominator separately, as

$$P = -\frac{\int d\lambda f(\lambda) Q(\lambda)}{\int d\lambda f(\lambda) I(\lambda)}, \tag{9.9}$$

where f is the spectral sensitivity of the simulated channel of the camera or sensor.

In general, BRF is a function of four angles, but if the target is sufficiently horizontally/azimuthally isotropic, the dependence is reduced to three, i.e., the only azimuthal variable is the difference ($|\phi - \phi_0|$). Newly fallen snow is generally isotropic, but older snow tends to have some anisotropy due to wind, sun, topography, and melting.

9.4 Instrumentation

The snow BRF measurements have been taken using many different instruments and setups by various research teams. One of the most advanced field goniometers is the Automated Spectro-Goniometer (ASG) at the University of California, Santa Barbara (Painter et al., 2003b), that is being used in many campaigns with impressive results. Of superior engineering is also the famous FIGOS at the University of Zürich (Odermatt et al., 2005; Schopfer, 2008), with possibly hundreds of measurements from various targets. A promising new setup is also the IAC ETH Goniospectrometer at ETH Zürich (Bourgeois et al., 2006b). All of these can be mounted and operated in snow. They use ASD Field Spec PRO FR or DUAL as a spectroradiometer, and a mechanism to move the optics around the target, although the principles of their mechanisms differ significantly from each other.

Of other setups we have no information, how operational they still are. Many measurements have been taken using rather ad hoc instruments or ones adapted from other purposes and probably cannibalized afterwards.

Two other kinds of instruments measuring the BRF of very large areas are the mast-based moving spectrometer at Dome C, Antarctica (Hudson et al., 2006), and the PARABOLA (Bruegge et al., 2000). Instead of rotating over the same spot, they scan over a large area, and have to assume that the surface is homogeneous over the full range.

The following subsections describe evolution of the Finnish Geodetic Institute goniometers in more detail.

9.4.1 Model 2, 1996: a simple one-angle manual field goniometer

Goniometry started at the Finnish Geodetic Institute using an ad hoc construction of abandoned metallic bookshelves with a borrowed spectrometer (model 1). It was never used for snow measurements, but gave many ideas for ensuring designs. A simple one-axis manual goniometer (model 2) was built by the Proto Centre of the Helsinki University of Technology for the first snow measurement expedition in 1996.

The goniometer had one $\pm 80°$ tilting arm and a triangular foldable frame on the ground. The arm could be tilted in 5° steps, manually. The azimuth was adjusted by moving the whole instrument manually around the target. Mounting of the system took about 20 minutes, and a measurement of one azimuth angle about 10 minutes, thus a full hemisphere took typically over 1 hour. The detector was a renovated Photo Research PR 713/PC spectrometer, controlled by an outdated 'portable' 286 Toshiba laptop with an ISA slot card. The system was calibrated using a 5-cm Spectralon white reference panel from Labsphere, which was hinged on the frame so that it could be turned to the right place and away very quickly. The reference measurements could thus be done relatively easily for all angles, although later it was concluded that taking references at every possible moment not only slowed down the measurement considerably, but did not actually improve the results, because the continuous up-down movement caused some sloppiness and delevelling of the standard.

In hard snow the operation was relatively easy. Deep and smooth snow made the operation complicated, the legs of the goniometer sank deeply into the snow, and moving the instrument was rather awkward. The frame corrupted the measurement area sometimes too easily, because the observation point was only 15 cm away from the frame. There was also some uncertainty about whether the nearby parts would cause unwanted shadowing, reflection, or heating of the target. Azimuthal movement always took more time than wanted.

The Photo Research spectrometer was not actually intended for winter use, and the shutter jammed every now and then, requiring some care. The computer screen needed a small tent around it to be readable in sunlight. The transportation of the instrument, using a van or sleigh, was easy.

The instrument gave important results of snow (Peltoniemi et al., 2005a), and good ideas for ensuring design.

9.4.2 Goniometer model 3, 1999–2005

To overcome the azimuthal movement problem, target disturbance and the small bending issues, a third goniometer (Peltoniemi et al., 2002; Widen, 2000; Peltoniemi et al., 2005a) was ordered from the Helsinki University of Technology. Model 3 consisted of two horizontal rings of a diameter of 2 m, one rotating above the other to change the azimuth angle, and a tiltable detector arch with an axle on the upper ring to change the zenith angle. All the movements were still manual, and the tilt was controlled using a winch. The system was levelled by 3 or 4 hydraulic legs. Mounting the system took about 1 hour, but then a full hemisphere was measured in 20 minutes. For snow measurement, the goniometer had optional skis, allowing fairly easy movement of the system for short distances without dismounting,

Fig. 9.2. Goniometer 3 being moved in Sodankylä 2001.

although it was a bit too heavy to move the system continuously (Fig. 9.2). For transport, a light trailer was prepared.

Initially, the goniometer was build for an old PR 713 spectrometer, and it was massive enough to carry the 5 kg weight without bending or shaking. However, an ASD Field Spec Pro FR spectrometer (350–2500 nm) was bought soon after the goniometer was ready. The ASD was equipped with a 4 meter long optical fibre connecting the fore optics to the main instrument, thus all the heavy parts could lie on the lower rotating parts of the system, and only the lightweight fore optics stressed the arch. Optics included 1°, 3°,8°, 18° and 27° opening angles, of which 1° and 3° were used mostly, making spot diameters of 3 and 9 cm at a distance of 2 meters. A laser pointer was installed parallel to the lens to locate the exact measurement point.

For calibration, a Spectralon white reference standard of $25 \times 25\,\mathrm{cm}^2$ was obtained. This allowed use of 3° optics, but was still too small for 8°. The Spectralon was usually levelled over the target using various ad hoc supports (stones, sticks, etc.) carefully so as not to disturb the target. Several more advanced mounting systems were designed, but no optimum design was found.

The measurement cycle was changed so that the white reference measurement was taken only from nadir, and then the target was measured from nadir to the lowest zenith angle and back as fast as possible. After that the reference was measured again, and if lightning conditions had remained stable, the result was accepted, otherwise abandoned and repeated. A diffuse light reflection measurement was taken also with a reference and target from nadir, and only occasionally from other zenith angles, since it was observed to be rather isotropic.

Goniometer 3 proved to be reliable, fast and easy to operate, once it was fully set up. Its main shortcoming was manual recording of measurement angles, which

generated a rather large amount of office work afterwards in typing and checking the data, and was too prone to writing, reading and typing errors. Other limiting factors were its heavy weight, requiring two strong men to move and mount it, its long mounting time and occasional shadows cast over the target by the arm and rings.

The new ASD Field Spec Pro FR spectrometer was a major improvement over the old PhotoResearch equipment. An issue to be noted especially with 1° optics, is that the spectrometer actually contains three separate sensors, each connected with a fiber bundle of its own, and these bundles each have different field of views, i.e. they see different parts of the target, which causes visible steps on the data when the target or illumination is heterogeneous (more on this and other difficulties later).

Goniometer 3 was successfully used for over 100 BRF measurements from snow (Peltoniemi et al., 2005a), vegetation (Peltoniemi et al., 2005b), aerial photogrammetry test field (Peltoniemi et al., 2007), and many other targets.

9.4.3 FIGIFIGO, 2005–

The experiences with the earlier goniometers guided in the design of a new very lightweight manual goniometer model 4, and an automatic motor-driven goniometer model 5. This evolved quickly into something known now as FIGIFIGO, or the Finnish Geodetic Institute Field Goniospectrometer, see Fig. 9.3. The FIGIFIGO system consists of a motor-driven moving arm that tilts up to ±90° from the vertical, variable fore optics in the high end of the arm, and an ASD FieldSpec Pro FR 350–2500 nm spectroradiometer. For laboratory measurements, the FIGIFIGO is installed on the rings of model 3, but in the field azimuth is changed again by manually sliding or carrying. Accurate angles are read with an inclinometer and an electronic compass, and most recently, an all-sky camera has been used to orient the system to the sun. The position is determined with GPS. The system can be mounted on a light sledge for snow measurement.

In addition to spectral BRF measurements, the FIGIFIGO can optionally measure linear polarization in the full spectral range. FIGIFIGO uses a calcite Glan–Thomson prism as a polarizer. The polarizer can be rotated 360° in 1° increments to observe different polarization directions. The first software version supported only 0° and 90° positions, suppressing the similarly interesting 45° and 135° polarizations (to yield Stokes U-parameter), but the newest version supports all states.

Typically, the footprint diameter is about 10 cm, elongating at larger sensor zenith angles as $1/\cos\theta$, and wandering around a few centimeters by bending and with azimuthal movements. A motorized fine-tune mirror has been recently installed to correct paraxial and bending errors and to keep the measurement point stable to an accuracy of 2 cm.

The instrument is calibrated as earlier by taking a nadir measurement from a Labsphere Spectralon white reference plate before and after each sequence. If the polarizer is used, separately for all polarizer directions. The Spectralon is carefully levelled at horizontal with a bubble level with an accuracy of about 1°.

The measurements are always started at 0° relative azimuth (principal plane). In the first measurement mode, the measurement arm is moved to selected measurement angles, and is stopped for the measurement. In the second mode, the arm

Fig. 9.3. The FIGIFIGO measuring BRF and polarization in Sodankylä in 2009. Optics is located horizontally in the upper end of the arm, and is looking down to the target through a mirror. The mirror has two small actuators to fine tune the spot location or scan over a small range. In front is the Spectralon white reference target, and behind the researcher's head is the sun and sky light monitoring pyranometer. [For a color version of this figure, see the color section.]

moves slowly from a 0° zenith angle to a maximum zenith angle (usually 70° or 80°) while the spectrometer is continuously recording data. Then the polarizer is turned 90° or 45°, and the arm turns back and continues to the maximum zenith angle in the other end, at which time the polarizer is turned again, and the arm returns upright or continues for another polarization direction. Then the instrument is turned to the next azimuth angle, and the sequence is repeated. Usually, 4 to 6 azimuth angles are measured, and if the weather allows, the principal plane is repeated as a final measurement. The spectrometer is calibrated at frequent intervals with the reference panel, depending on assumed and observed stability of the illumination, after 1–3 azimuth turns.

The diffuse skylight is measured at the beginning of each sequence from nadir by blocking the direct sunlight with the help of a small screen (diameter of 40–60 cm) at a distance of 2–4 m. Measurement of one hemisphere at one illumination angle takes between 15 and 60 minutes, depending on the sky condition, expected results and accuracy, and various technical aspects.

FIGIFIGO is still at the time of writing (summer 2009) in full use, with hundreds of measurements (Peltoniemi et al., 2009; Suomalainen et al., 2009b,a; Puttonen et al., 20099, and subject to continuous improvements.

9.4.4 Light sources

For many measurements, the sun is the ideal choice of light, but it is not without problems. The availability of the sunlight is limited and unpredictable. Thus one often needs to wait days or weeks for adequate weather. Even on clear days, the atmosphere varies in transmittance and scattering, causing significant variations for the incident light, in all time scales from seconds to days. The sun is continuously moving in the sky. Thus, one needs to take reference measurements as often as conveniently possible (1–5 minutes), or monitor the incident light parallel with reflectance measurements. Most ideal is to use identical second spectrometer monitoring the incident light (Odermatt et al., 2005; Bourgeois et al., 2006b). FGI has been using a Kipp&Zonen SP lite silicon pyranometer since 2006, and taking white reference measurements based on the smoothness of the pyranometer curve. Before 2006 sky clearness was observed only visually and reference cycle was kept as short as possible.

A further problem is that the Sun cannot be set to a selected angle, but the only way to change illumination direction is to wait until the sun moves to the desired angle. Unfortunately the conditions rarely allow multiple measurements of the same unchanged target several times of the day. Sunlight is also rather hot, and accelerates the melting and metamorphosis of the snow. Thus, use of artificial light sources that can be set to selectable positions and angles will extend the measurement possibilities significantly.

In making a good artificial light source for spectral BRF measurements, there are 5 major challenges to be met:

(1) Spectral sensors need enough light power for a good signal to noise ratio.
(2) The light source must be stable over the short internal scanning cycle and the whole measurement period.
(3) The light source must have reasonably smooth and flat spectrum, extending over the observation range without spikes or dark bands.
(4) For a consistent BRF measurement, the light pattern must be homogeneous and collimated throughout the measurement area (for internally transmitting targets, such as snow, even larger).
(5) The light must not polarize unwantedly.

Most laboratory light sources were found to be far too underpowered, and industrial and entertainment lights too unstable. A 300-W halogen was tried in the measurements 1997 and a 500-W theatre light in the 2000–2004 measurements with of partial success. A rather long integration time ($\gg$1 s) was needed to improve the signal and smooth the 50-Hz AC pattern.

A much better solution was found with the Oriel 1000 W Quartz Tungsten Halogen (QTH) light source, providing very stable and strong light over the important spectral range.

The challenge of homogeneous light pattern appeared most tricky. Available light bulbs are not ideal small point sources nor even flat surface sources, but have

a filament pattern. With enough energy to waste, one can of course create very homogeneous illumination by positioning a single light bulb far enough away to overcome the $1/r^2$ effect between the front and the back side of the spot, but in practice this solution gives far too little light. A better solution is a small-sized light bulb and a large collimating lens, about the same size as the desired spot. Such lenses of a diameter of 50 cm or more are hard to find, and plastic Fresnel lenses do not transmit infrared light well. Thus, an attempt was made to correct the light pattern using the largest available glass lens of a diameter of 20 cm and focal length of about 60 cm. The lens was connected to the lamp using tubes and a 90° Oriel mirror, that allowed easy turning of the illumination angle. The lamp was then mounted on a telescopic 180–400 cm tripod.

Fig. 9.4. Night measurement using Oriel lamp and old Goniometer 3 in Sodankylä. Sometimes there are spectacular moments in the life of field researchers.

The second invention was to use a polished aluminium off-axis parabolic mirror with a diameter of 50 cm. The mirror collimates light and turns the beam up. Another flat mirror in an adjustable mount turns the beam down to the sample, in selectable angles. This setup yields a rather clean spot of 60–100 cm. Even with this, the inhomogeneity of the light beam was 10% to 20% within the measurement area, due to imperfections in the mirrors and lamp filament pattern.

9.4.5 Data processing

Data processing procedures evolved over the years, the automation of measurements contributing to the development. In the manual goniometer era, the first step was to type the angles into a computer-readable file, laboriously checking writing, reading, and typing errors, and other problems in the data. In more modern times, this step was simplified to visual and semi-automatic checks for the consistency of data, locating outliers, discontinuities, and other evident errors. Data that was clearly

wrong was abandoned and data that was questionable or imcomprehensible was marked so that it was not automatically read in any further processing, but nevertheless remained available, in case a physical reason for some strange behaviours is later found.

The measured unnormalized radiance spectra S was normaliszed by measured nadir spectra from a reference target (S_{STD})

$$R(\mu, \mu_0, \phi, \phi_0) = \frac{S}{S_{\mathrm{STD}}} R_{\mathrm{STD}}, \tag{9.10}$$

where R_{STD} is the reflectance of the reference target. Ideally $R_{\mathrm{STD}} = 1$, but all real materials differ somewhat. Here, a form $R_{\mathrm{STD}} = R_0(\lambda, 0°, 60°) r(\mu, \mu_0)/r(0°, 60°)$ is assumed, where $R_0(\lambda, 0°, 60°)$ is the spectral reflectance given by the Spectralon manufacturer, in one geometry, and r is wavelength independent angular part, approximated using various published (Courréges-Lacoste et al., 2003) and unpublished data. When possible, the reference spectrum was measured twice, before and after each sequence, and if slightly different, e.g., due to rising sun or changing atmosphere, a linear interpolation was used. When diffuse light was present, the diffuse background (S^{DIFF}) was subtracted from the measured spectra as

$$R(\mu, \mu_0, \phi, \phi_0) = \frac{S - S^{\mathrm{DIFF}}}{S_{\mathrm{STD}} - S^{\mathrm{DIFF}}_{\mathrm{STD}}} R_{\mathrm{STD}}, \tag{9.11}$$

where $S^{\mathrm{DIFF}}_{\mathrm{STD}}$ is the reference standard measured in the diffuse light only. For polarized light, the measured unnormalized radiance spectra S_i for each polarization state $i = [h, v, \backslash, /]$, horizontal, vertical, 45°, and 135°, was normalized

$$R_i(\mu, \mu_0, \phi, \phi_0) = \frac{S_i}{S_{\mathrm{STDh}} + S_{\mathrm{STDv}}} R_{\mathrm{STD}}, \tag{9.12}$$

or with diffuse light correction

$$R_i(\mu, \mu_0, \phi, \phi_0) = \frac{S_i - S_i^{\mathrm{DIFF}}}{S_{\mathrm{STDh}} + S_{\mathrm{STDv}} - S^{\mathrm{DIFF}}_{\mathrm{STDh}} - S^{\mathrm{DIFF}}_{\mathrm{STDv}}} R_{\mathrm{STD}}. \tag{9.13}$$

Note that because Spectralon is polarizing (Suomalainen et al., 2009a; Haner et al., 1999; Georgiev and Butler, 2004), $S_{\mathrm{STDh}} \neq S_{\mathrm{STDv}}$.

The analysis was simplified above by the observation that in all measurements the reflected diffuse light was found isotropic within 20%, typically below 10%. Since the diffuse contribution is in all acceptable conditions below 20% (mostly below 10%), the errors from anisotropy remains minimal for end results, and it was assumed in further data processing that the diffuse component is perfectly isotropic.

In later measurements, where the incident light was monitored using a pyranometer, its readings were further used to fine-tune the calibration. If the incident light variations were within the range 1–5%, the measured spectra were normalized using the incident irradiance value from the pyranometer E as $S' = SE_0/E$, where E_0 is the value taken at the start. If the variation were even larger, the measurement was aborted and the data abandoned, because one could no longer rely on the variation being consistent throughout the spectrum.

The Stokes I, Q and U parameters (in effect the Muller matrix elements R_{11}, R_{21}, and R_{31}) were computed for each measurement point as $R_{11} = R_h + R_v$, $R_{21} = R_h - R_v$, $R_{31} = R_\backslash - R_/$. Because the different polarizations were measured in their own sweeps, the zenith angles did not match, and linear nearest-neighbour interpolations was used. During the 2007 measurements, only two polarization states were supported (h and v, or I and Q), but the current configuration allows now measurements of 45° and 135° polarization states, too, providing data for the Stokes U parameter.

The measurements and subsequent analysis and plots are mostly made under the assumption of left–right symmetry, which halved the needed measurement time in the field. The justification for this assumption was checked every now and then, and found to apply in the range of 5% to 20%, i.e., not automatically adequate.

The accuracy of the spectral BRF measurements using FIGIFIGO is estimated to be 2–3% in the visible band and a good conditions, with a polarization accuracy of 2–5%. Suboptimal weather, which begins with hardly visible clouds, can decrease accuracy quickly by more than 10%. Angle registration accuracy is 2°.

9.5 Main research efforts

A number of measurement efforts have been run between 1997 and 2009, as shown in Table 9.1 and in Fig. 9.5. In practice, all of them have been a side product of some other campaign, project or purpose, but they gave a good opportunities to study snow reflectances on their own.

Snow measurements started 1998 with a particular research effort in Vuotso, north Finland, using the second field goniometer of FGI (Peltoniemi et al., 2005a). In the beginning weather was clear and cold. The snow was days or weeks old powder snow, forming a loose layer with a thickness of about 1 m, requiring skis or snow shoes to be used for human transport. This snow was measured in several locations around the area in sunlight. Then the weather to changed cloudy and warm, and the wet snow was measured over two nights, using a 300-W halogen lamp. Many interesting observations were made, but the diversity of the data was so puzzling, that no strong conclusions could be drawn.

Measurements continued using the third goniometer between 2001 and 2004, mostly in Sodankylä, yielding large amounts of useful data from several dry and wet snow types (Peltoniemi et al., 2005a). Various artificial snow processing techniques, collecting to piles, compressing, flattening, roughening, digging from different levels, etc. were tried. The results gave an excellent overview of snow reflectance properties, and plenty of ideas for further research. Nevertheless, the sampling remained still too sparse and unsystematic for many reproducible conclusions with which to answer many interesting questions. It became clear that a faster and more automatic system was needed to get more systematic and coherent data.

A significant improvement to measurement techniques came in 2005 with the new FIGIFIGO. Its faster setup and increased measurement rate more than tripled the amount of daily data, and automatic recording and an auxiliary pyranometer improved quality. A more advanced light source setup made artificial light measurement feasible and more reliable.

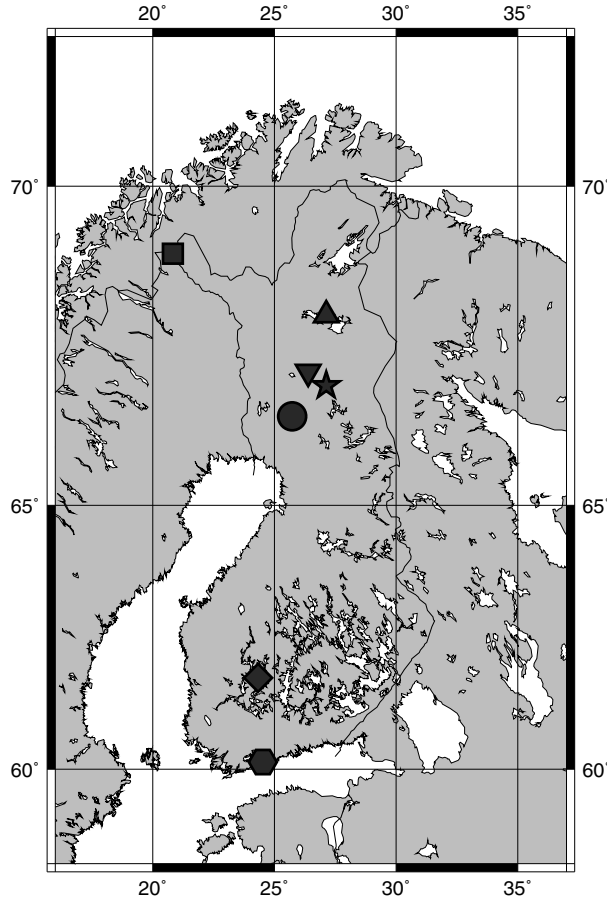


Fig. 9.5. Locations of the measurement efforts, Vuotso (triangle), Hyytiälä (diamond), Kilpisjärvi (square), Rovaniemi (circle), Pyhätunturi (star), Masala (hexagon), and Tähtelä, Sodankylä (inverted triangle).

A quick set of laboratory measurements was made during the winter of 2005 in Masala. The sample was prepared outside and taken inside the goniometer at room temperature. The sample was continuously measured during the melting process, and allowed to freeze again outside (1 hour to overnight). This yielded controlled and semi-reproducible results, but the melting process appeared to be too radical to simulate natural melting, with evident condensation of air humidity on snow. The small sample size (30 cm × 50 cm × 10 cm) also caused clear effects, for example the red colour of the container was clearly visible through the 10 cm snow layer.

A new expedition was run in Sodankylä in April 2005, using both sunlight and the Oriel QTH source. The snow was mostly days or weeks old, varying between dry and moist. Surface micro-topographies were rather flat. Samples were selected in various open or semi-open locations near the research station, and once near Pyhätunturi fell. Wetness was measured using the Toikka snow fork, but its resolution of about 5 cm in depth was found to be too coarse for detailed snow surface

Table 9.1. Snow measurements expeditions using goniometers 2 and 3 with a short summary of the target. The columns indicate: measurement date, location, setup (G2 = goniometer model 2, G3 = goniometer model 3, H = 300-W outdoors halogen lamp, L = Oriel 1000-W lamp), angles of incidence in the measurement order (↑ = only principal plane, † = partial BRF, ∗ = full quasisphere (assumes left–right symmetry), ∗full = true hemisphere), wetness of the sample (D = dry, W = wet, number = measured wetness [%], f = freezing, m = melting), grain shape classification according to Colbeck et al. (1985), and some other relevant note.

Date	Location	Setup	Angles of inc.	Wetness	Snow-type	Note
16.4.1998	Vuotso 1	G2	58°†	D	sr	
17.4.1998	Vuotso 1	G2	60°∗	D	sr	
18.4.1998	Vuotso 2	G2	60°∗	D	sr	surface 15° tilted
20.4.1998	Vuotso 3	G2H	60°	Wf	mf	
21.4.1998	Vuotso 3	G2H		W	mf	
23.4.1998	Vuotso 1	G2		W	mf	
7.3.2001	Hyytiälä	G3	67†	D	nd,dc	
29.3.2001	Sjökulla	G3	55°†	M	sr	10 cm
3.5.2001	Sodankylä 1	G3	60°†	W	mf	
4.5.2001	Sodankylä 1	G3	65°∗ –51°∗	M–W	mf,sl	
5.5.2001	Sodankylä 1	G3	61°∗–53°∗	W	mf	dirty
8.5.2001	Sodankylä 1	G3	50°∗	W	mf	dirty, manipulated, 8 cm
7.3.2002	Rovaniemi	G3	72°↑	D	sd	
8.3.2002	Rovaniemi	G3	72°↑	D	sd,dc	
10.3.2004	Sodankylä 3	G3L	55°∗,52°∗	W	dc	
11.3.2004	Sodankylä 3	G3	71°–72°∗	M	lr	
11.3.2004	Sodankylä 3	G3L	49°∗,44°∗,62°∗	Mf	lr	
13.3.2004	Sodankylä 3	G3L	62°∗,52°∗,45°∗	D	dc	
13.3.2004	Sodankylä 3	G3L	45°∗,68°∗	D	mx	surface broken
14.3.2004	Sodankylä 3	G3L	6°5,6°1,48°,43°	D	dc	
14.3.2004	Sodankylä 3	G3L	45°,60°	D		compressed
16.3.2004	Sodankylä 3	G3L	57°∗,50°∗,65°∗	D	lr	
10.2. 2005	Masala, lab	G3L	39°↑,60°↑↑	M	lr,cf,mf	

measurements, as in places the wetness profile was found to vary in the scale of 1 cm by visual and manual inspection. The grain size and shape were measured both visually and photographically.

The next research effort was again run in Tähtelä Research station, Sodankylä, in April 2007. The targets were selected in a semi-open area near a newly mounted Norsen mast. All measurements were made in sunlight, with relatively clear weather. The snow was melting several centimeters a day, but freezing again during the night. The snow was still relatively pure, with some debris from trees and the environment here and there. Its structure was rather rough, with about 4 cm height variations in the scale of 20 cm. Many repetitions were made, to study the reproducibility of results and local heterogeneity. As a novel aspect, during the

Table 9.2. Snow measurements expeditions using FIGIFIGO, as in Table 9.1. The locations are Sodankylä 2 = front yard of the guest house at Tähtelä Research Centre in Sodankylä, Sodankylä 3 = small opening behind the guest house at Tähtelä Research Centre in Sodankylä, Sodankylä S = a small open swamp in a forest near Tähtelä Research Centre in Sodankylä, Sodankylä J = Jänkävuopajanaapa, big open swamp 1 km from Tähtelä, Sodankylä N = near the Norsen mast at Tähtelä, Sodankylä), setup (FU = FIGIFIGO unpolarized, FP = FIGIFIGO with a polarizer, L = Oriel 1000-W lamp).

Date	Location	Setup	Angles of inc.	Wetness	Snow-type	Note
17.4. 2005	Sodankylä 2	G5	54.0°–52.4°	D	sr	rough
17.4. 2005	Sodankylä 3	G5	72.0°–75.6°	D	lr	rough
18.4. 2005	Sodankylä 3	G5L	61.9°, 55.0°	D	sr	
19.4. 2005	Sodankylä 3	G5	62.0°–58.8°∗	D	sr	
19.4. 2005	Sodankylä 3	G5L	55°∗, 64°∗, 45°∗	Mf	sr	
20.4. 2005	Sodankylä S	G5	61.9°–58.8°↑∗∗	Dm	lr	
21.4. 2005	Sodankylä J	G5	75.8°–72.2°†	D	lr	
22.4. 2005	Sodankylä J	G5	55°↑	D	lr	
22.4. 2005	Sodankylä J	G5L	45°∗, 62°∗	W	cl	
23.4. 2005	Pyhätunturi	G5	62.7°–56.3°∗full∗full	Dm	cl	
17. 4. 2007	Sodankylä N	FU	61.1°–59.6°∗∗	D	lr	
17. 4. 2007	Sodankylä N	FU	57.4°–58.5°∗∗	W	lr	
18. 4. 2007	Sodankylä N	FU	70.7°–59.0°↑∗↑↑	Dm	lr	
18. 4. 2007	Sodankylä N	FU	57.4°–62.0°	Wm	lr	
20. 4. 2007	Sodankylä N	FP	70.8°–55.9↑↑↑∗↑↑ ↑°	W	lr	
1. 4. 2008	Sodankylä N	FUL	43°∗,62°∗,55°∗,71°∗	D	9lr,1mx	
2. 4. 2008	Sodankylä N	FPL	66°∗,37°∗,45°∗	D	9lr,1sf	
3. 4. 2008	Sodankylä N	FP	70°∗∗–62°∗∗	D	9lr,1sf	
4. 4. 2008	Sodankylä N	FPL	35°∗,55°∗	D	9sd,1dc	
5. 4. 2008	Sodankylä N	FPL	42°∗,55°∗,65°∗	D	9nd,1dc	
19.3.2009	Masala	FP	66–63°∗∗	D	sr	
20.3.2009	laboratory	FPL	60°∗∗∗, 45°∗, 72°∗	DW	lr	20 cm
20.4. 2009	Kommattivaara	FP	57°↑↑↑	W	lr	M
21.4. 2009	Sodankylä N	FPL				
22.4. 2009	Mantovaara-naapa S	FP	65°∗	M	sr,cl	smooth
22.4. 2009	Mantovaara-naapa R	FP	60–55°↑∗∗↑↑↑	M	mf	rough
23.4. 2009	Korppiaapa R	FP	65°∗	W	mf	rough
23.4. 2009	Korppiaapa I	FP	60°∗	D	bi	rough ice

last day the reflectance was measured using a calcite broadband polarizer both in horizontal and vertical polarization states.

The measurements continued again in the spring of 2008 at the same place, as part of a larger Snortex campaign. The weather was cloudy or semi-cloudy for most of the days, so that only one daylight measurement was possible; otherwise measurements were taken during nights with the Oriel lamp at 2–4 illumination

Table 9.3. An incomplete extraction of other known snow BRF measurements.

Reference	Date	Location	Setup	Angles of inc.	Snow types	Notes
Middleton and Mungall (1952)	1951–52	Ottawa	portable gonio-photometer	0°, 30°, 45°, 60°, 75°	many	
Steffen (1987)	Jul–Aug 1986	Urumuqi Glacier, China	manual, 500–600 nm	28°–76°	powder, new, old	
Warren et al. (1998)	Feb 1986	Dome C, East Antarctic Plateau	900 nm sensor in mast	75°	dry, rough	roughness effects
Warren et al. (1998)	Jan–Feb 1991	Dome C, East Antarctic Plateau	900 nm sensor in mast	67°,72°, 73.5°	dry, rough	roughness effects
Warren et al. (1998)	1992	Dome C, East Antarctic Plateau	600 nm sensor in mast	67° – 90°	dry, rough	roughness effects
Warren et al. (1998)	Dec 1992	Dome C, East Antarctic Plateau	900 nm sensor in mast	67° – 90°	dry, rough	roughness effects
Hudson et al. (2006)	1996–2006	Dome C, East Antarctic Plateau	ASD in mast	51° – 87°	dry, rough	roughness effects
Nolin and Dozier (1991)	May 1989	Mammoth Mountain, Sierra Nevada	SE-590		thin	
Nolin et al. (1994)		Mammoth Mountain, Sierra Nevada	SE-590		moist	
Nolin et al. (1994)		CU/ETH station, Greenland	GER	65°		
Leroux et al. (1998)	Mar 1995	French Alps				grain shapes, polarization
Aoki et al. (2000)	–1999	Hokkaido	ASD	56°		
Tanikawa et al. (2006)	Feb 2001	Hokkaido	ASD in tripod arm	55°	sd	
Tanikawa et al. (2006)	Mar 2001	Hokkaido	same	45°	granular, crust	
Painter and Dozier (2004)	Feb 2001	Mammoth lake, Sierra Nevada, Ca	ASG	47°	dc	
Painter and Dozier (2004)	Mar 2001	Mammoth lake, Sierra Nevada, Ca	ASG	47°, 41°	mf	
Nolin and Stroeve (2001)	Feb–Mar 2001	Yampa Valley, Colorado	PARABOLA			comparisons to MISR, AirMISR
Hendriks et al. (2004b,a)	Mar 2003	Davos	Utrecht goniometer	48°–73°	several	
Hendriks et al. (2004b,a)	Jun 2003	Morteratsch glacier, Switzerland	Utrecht goniometer		ice	
Kneubühler et al. (2008)	2007-08	Davos	FIGOS			

angles. The first of the measurements was taken unpolarized, the next four with polarization. The snow was relatively dry all the time.

In the winter of 2009 one set of measurements was taken in a field near Masala under perfect sunshine. The snow was dry and relatively smooth, clean and homogeneous. Air temperature was about −5°C.

Next day the melting measurement in the laboratory was repeated. Now the lab was allowed to cool for two hours by keeping doors open and ventilation at full speed. Outside temperature was around −5°C, inside +5°. Initially, the snow was rather similar to previous field samples. A snow sample of 20 cm thickness was taken from outside, and the measurement was started immediately, measuring continuously at 60° illumination zenith angle until the snow was considered very wet. Then lamp was moved to 45° and 72° angles and the BRF of wet snow measured in wider illumination range.

The Snortex campaign continued in Sodankylä March to May 2009, and further snow reflectance measurements were taken during the campaign in April around the area. Extensive snow measurements using various other instruments were taken at the same locations. The first point was in a sparse forest of short trees behind the hill of Kommattivaara. Snow was starting to melt in sunlight. Next night, freezing snow was measured near the Norsen mast using a lamp at two angles of incidence in four polarizer directions. An extensive set of measurements was taken in Mantovaaranaapa swamp, supporting several airborne cameras. The snow had been melting about two weeks, and in the morning the snow surface was very hard. Small amounts of fresh snow had fallen in previous days, and wind had collected this to random spots of a width of about 20 to 50 cm, and depth of 1–2 cm. This snow had transformed to rounded mid-size grains, and was still soft and smooth on the surface. The underlying hard snow was very rough. First, the smooth snow was measured, then the rough snow in three locations, to get a better sampling. In the beginning, the snow was relatively dry after a cold night, but started rapidly melting. In the afternoons it could not carry even a light sledge, although in the mornings easily carrying a man. During the last day, another rough melting snow sample as well as a patch of bare, rough ice was measured in Korppiaapa swamp. All the snow samples in the swamps were rather shallow, with a depth from 10 to 20 cm. A deeper analysis of the snow properties and the results of the research effort will be published in some designated campaign publications

9.6 Modeling

The snow layer is modeled as a random packing of randomly deformed and oriented ellipsoidal or rectangular ice grains (Peltoniemi et al., 1989; Peltoniemi and Lumme, 1992; Muinonen et al., 1996), with an optional liquid water coating, and a spectrally dependent complex refractive index (Warren et al., 1995). The scattering is solved using the Jray Monte Carlo ray–tracing program by Peltoniemi (2007). The program solves for the 4×4 BRF Muller matrix in the spectral range of 350 to 2500 nm.

In short, the algorithm is

(1) shoot a ray with a spectrum of unit Muller matrices from an initial position to an initial direction (optionally random)
(2) trace the ray until it hits the border of the scattering medium, and then see how it intersects various subvolumes inside
(3) inside a scattering subvolume, trace whether the ray hits the surrounding of a stochastic particle
(4) locate the random scattering point inside the stochastically scattering volume
(5) register the travelled path length
(6) compute the scattering to each preselected detector
    (a) angles of incidence and emergence, required surface normal, probability density function for this normal orientation
    (b) Fresnel coefficients, rotations
    (c) propagation probability to the sensor
    (d) record observation
(7) generate random normal
(8) compute Fresnel matrices and rotations
(9) select between a reflected or refracted ray
(10) continue tracing from point 2 until the ray escapes the medium
(11) record the ray
(12) compute absorbed energy spectrally using a wavelength-dependent absorption coefficient and recorded path length in each submedia

The subtleties include spectral weighting of the ray components during wavelength-dependent scattering, thin layered surfaces, free positioning of the sensors in the most interesting directions, e.g., backward, forward, specular, Brewster angles. Combining ray-tracing with adding/doubling and other accelerating techniques is under continuous construction.

Diffraction is totally ignored. Not only is it rather immaterial whether some parts of rays are slightly bent when passing a grain or considered to be unscattered, but in a densely packed medium the assumptions of the simple diffraction model fail: diffraction by what? – a grain, a hole between grain, semi-infinite surface, near or far field? Actually, diffraction models fail because of forward interference which happens in many other radiative transfer cases too and which it is better to ignore than to use an inappropriate model. Note also, that ray optics belongs completely to classical physics; thus, in the theoretical sense, a single ray contains infinite number of photons, and the common trend to talk about tracing individual photons is completely wrong and misleading.

Three different sample cases have been computed:

(1) Fine-grained new snow
    – 90% rectangulars of semi-axes 0.1 mm, 0.1 mm, 0.5 mm
    – grain surface roughness parameters $\sigma = 0.02$ and $\rho = 0.5$
    – 10% rounded grains, semi-axes 0.3, 0.4, 0.5 mm
    – packing density 10% (occupied volume/total volume), overlap allowed
    – flat snow layer
    – thickness 45 cm
(2) Old dry snow

- 90% rounded grains, grain semi-axes 0.5 mm, 0.75 mm 1.0 mm
- 10% rectangulars of semi-axes 0.3 mm, 0.4 mm, 0.5 mm
- grain surface roughness parameters $\sigma = 0.02$ and $\rho = 0.5$
- packing density 40% (occupied volume/total volume), overlap allowed
- optional water coating thickness 0.02 mm
- snow surface auto-correlation length 1 cm and height standard deviation 0.25 cm
- thickness 45 cm

(3) Wet snow
- grain semi-axes 0.5 mm, 0.75 mm 1.0 mm
- grain surface roughness parameters $\sigma = 0.02$ and $\rho = 0.5$
- packing density 50% (occupied volume/total volume), overlap allowed
- 5% of grains liquid water droplets
- snow surface auto-correlation length 1 cm and height standard deviation 0.25 cm
- thickness 45 cm

9.7 Results

Before going to real measurement results, the predictions by simulations are presented in Figs 9.6 to 9.9. All cases show a clear forward scattering spike, being darkest around nadir or a little forward, and with a very weak enhancement backwards. The spectrum is very sensitive to grain size. Grain shape affects forward and backward scattering somewhat. Regular crystals make the forward spike narrower, and increasing irregularity increases backward scattering. Polarization is strongest near the Brewster angle when the absorption is strong enough that single surface reflection dominates. The 45/135° polarization (Stokes U parameter) and circular polarization (V parameter) are low and almost insignificant compared to Q polarization. Polarization seems to depend on grain size via absorption: small grains, small absorption, more multiple scattering, less polarization. The differences between wet and dry snow seems minimal: only a small spectral effect (more about that later) but there are no directional or polarizational signals at all.

Next, the measurement results are shown and discussed (Figs 9.10 to 9.28). The general observation is that each sample is individual. Because of the metamorphosis, the sample often changes, even during the measurement, so that one cannot always be sure what was measured. Also, no parameter varies alone, but the whole snow in all of its properties, in all depth scales. Thus, drawing firm conclusions from the data is difficult.

From published results it is already known that different snow types differ by anisotropic, spectral and polarization properties (Painter and Dozier, 2004; Hudson et al., 2006; Nolin et al., 1994; Peltoniemi et al., 2005a, 2009; and many others). Here much more data is shown. First, an overview of snow BRF, and snow's spectral and polarized properties are shown. Then the clearest differences between snow types are pointed out, and more details specifically related to melting and wetness follow.

Snow BRF and spectral curves averaged over all measured samples are shown in Figs 9.10 and 9.11. Snow is a strong forward scatterer, but not a specular scatterer;

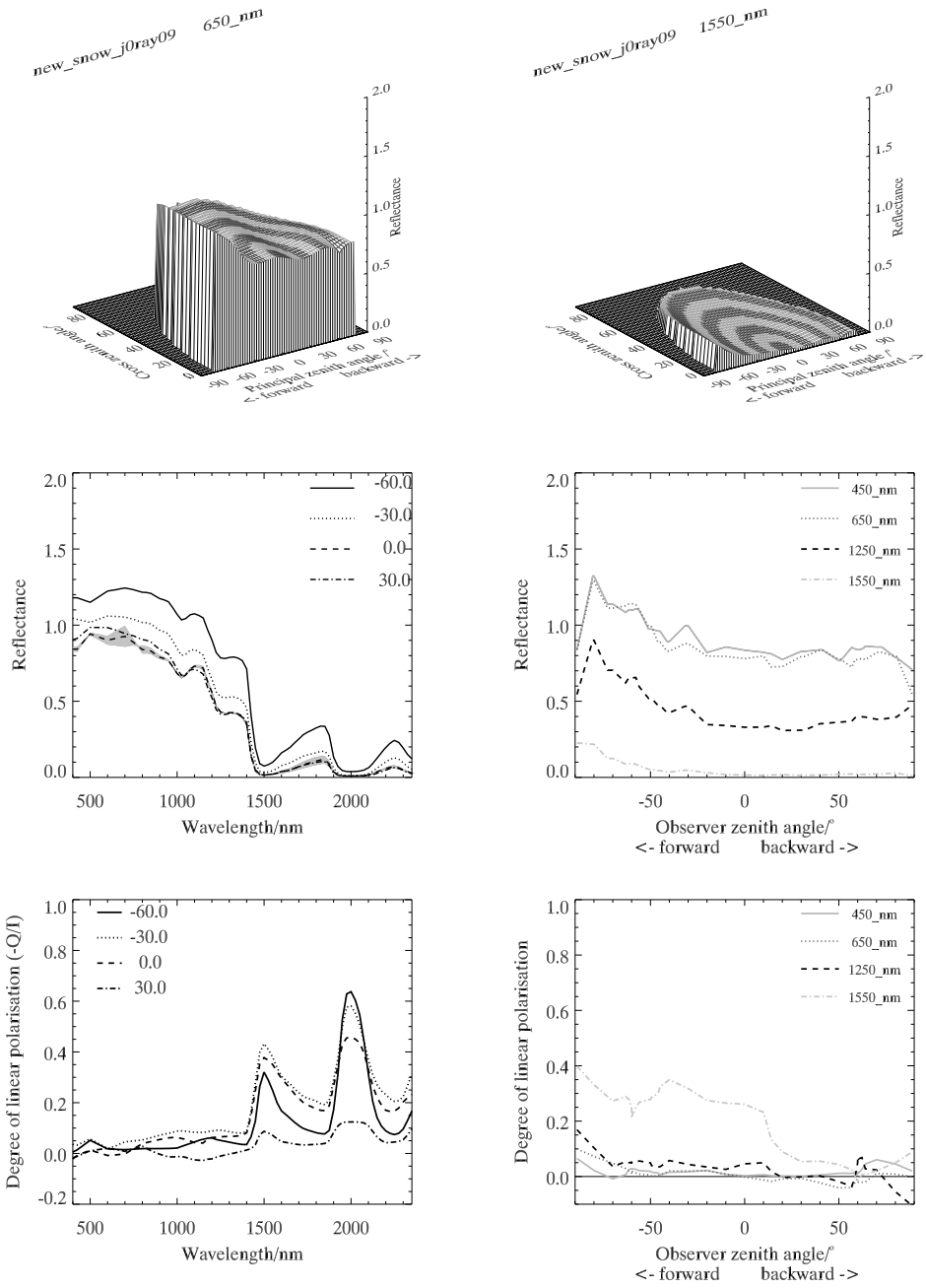


Fig. 9.6. Simulated BRFs of new snow, from top left: BRF cake in 650 and 1550 nm, the reflection spectra in four directions on the principal plane (colour indicates difference between two sampling), the BRF on the principal plane in four channels, the polarization spectra in four directions on the principal plane, and the polarization curve on the principal plane in four channels. The angle of incidence is 60°. Minor oscillations in the data are Monte Carlo noise. [For a color version of this figure, see the color section.]

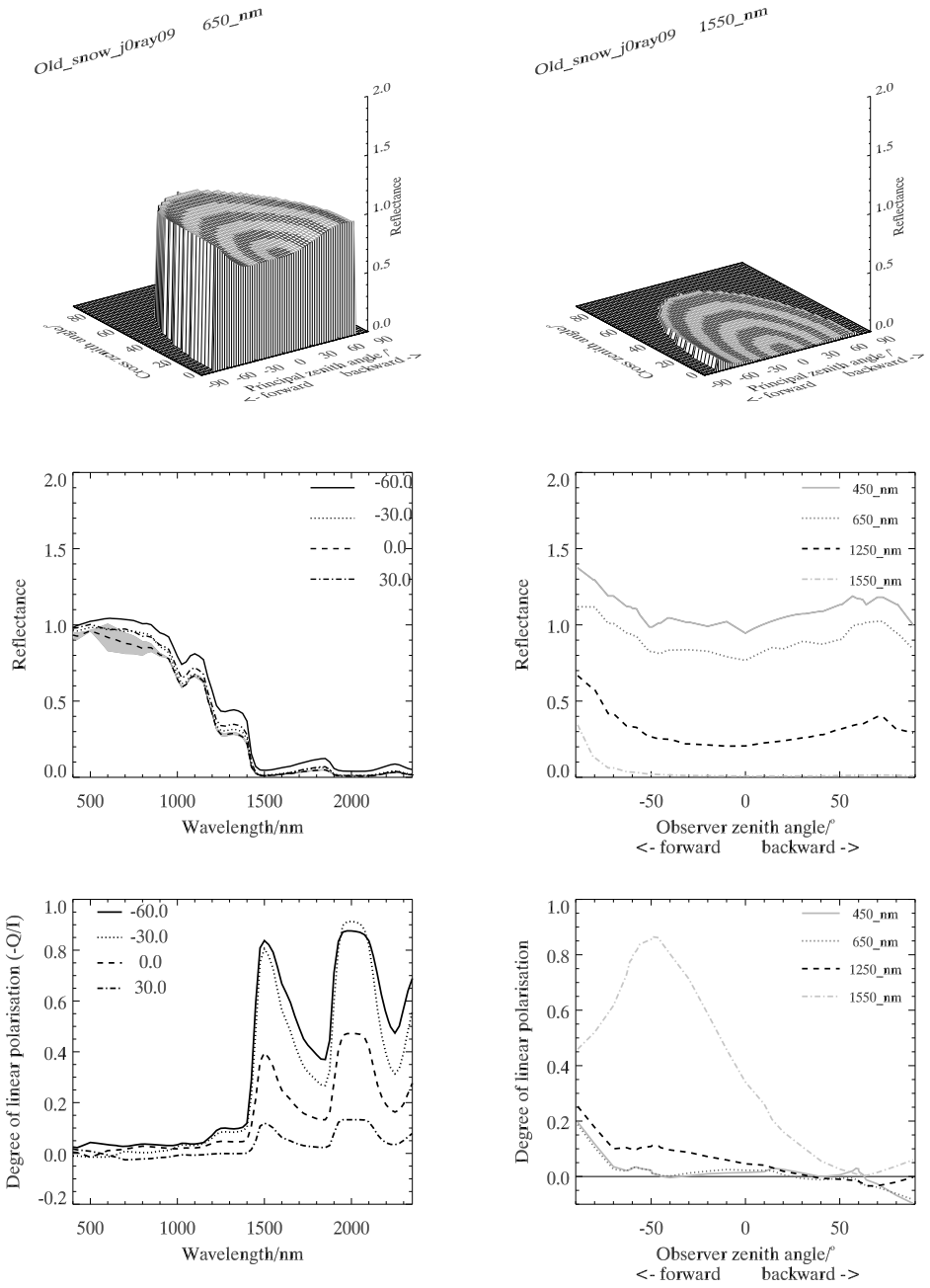


Fig. 9.7. Simulated BRFs of dry snow, as in Fig. 9.6. Note that the old snow spectrum is darker in infrared than in new snow, and polarization gets stronger. BRF shape changes are small. [For a color version of this figure, see the color section.]

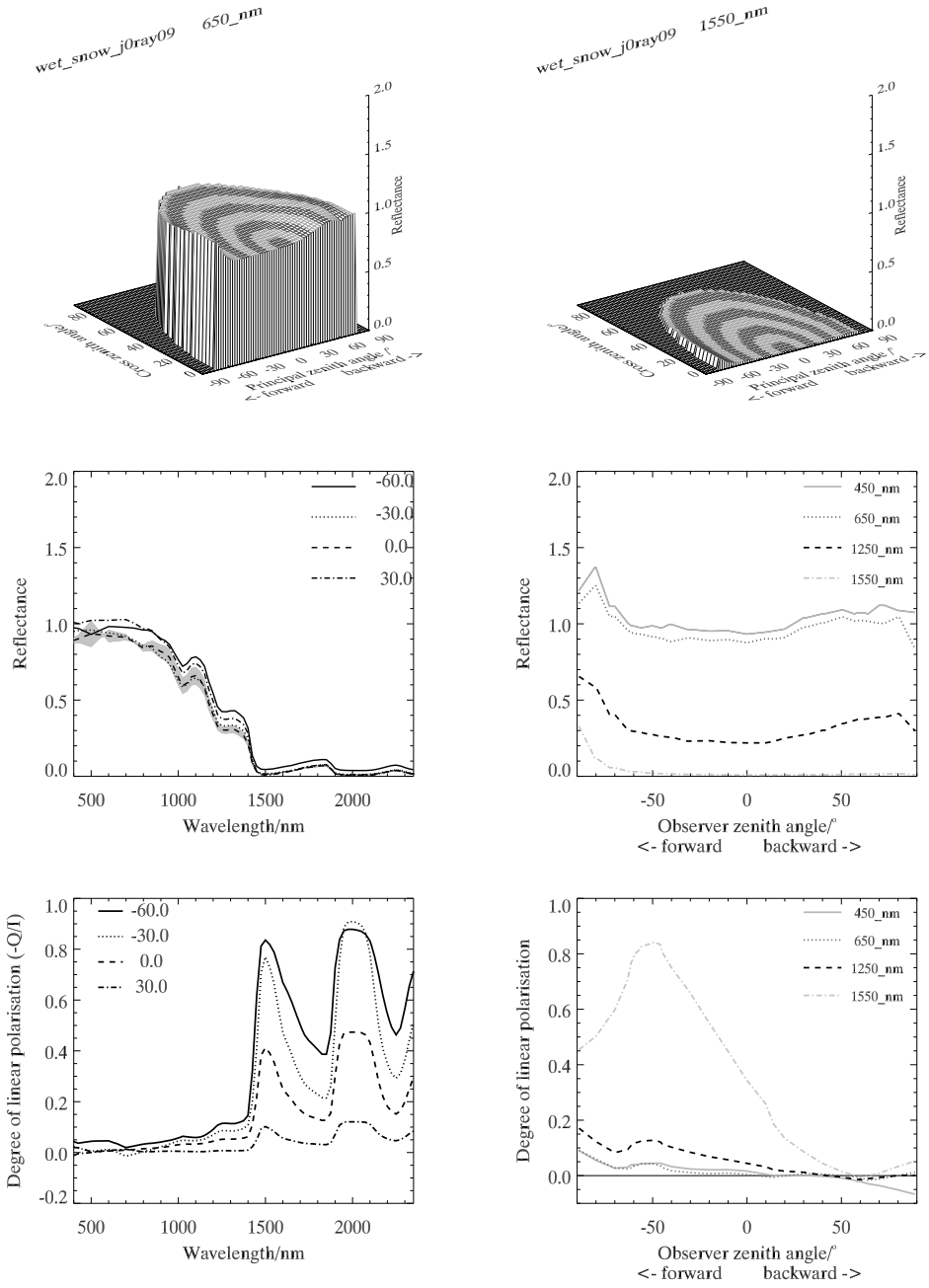


Fig. 9.8. Simulated BRFs of wet snow, as in Fig. 9.6. The only significant difference from dry snow is the very small spectral changes in some wavelengths, to be discussed in the next subsections. [For a color version of this figure, see the color section.]

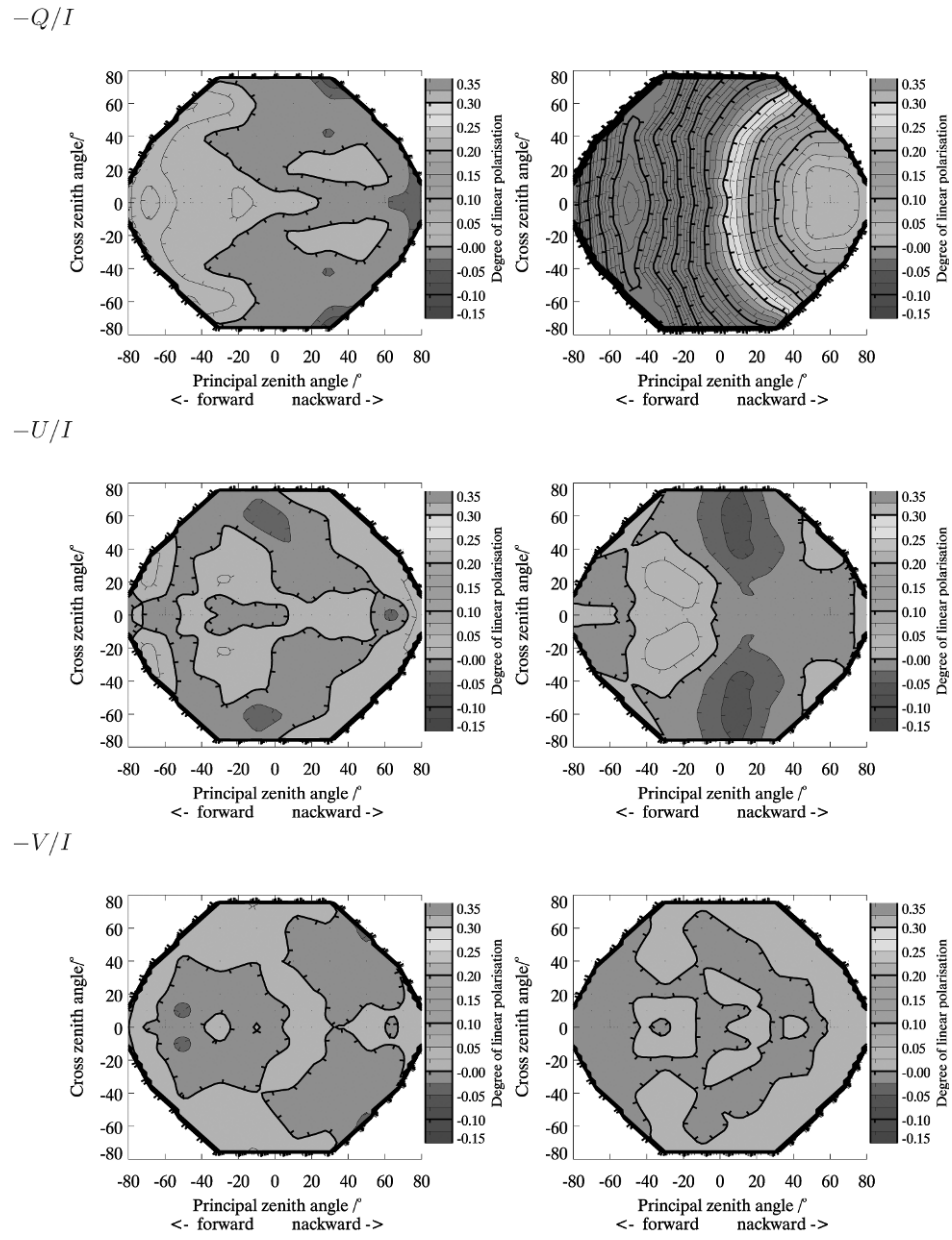


Fig. 9.9. Flat polarization plots of simulated dry snow, on the left 550 nm, on the right 1550 nm. From top to bottom $-Q/I$, $-U/I$, and $-V/I$. Polarisation is here defined in the scattering plane. [For a color version of this figure, see the color section.]

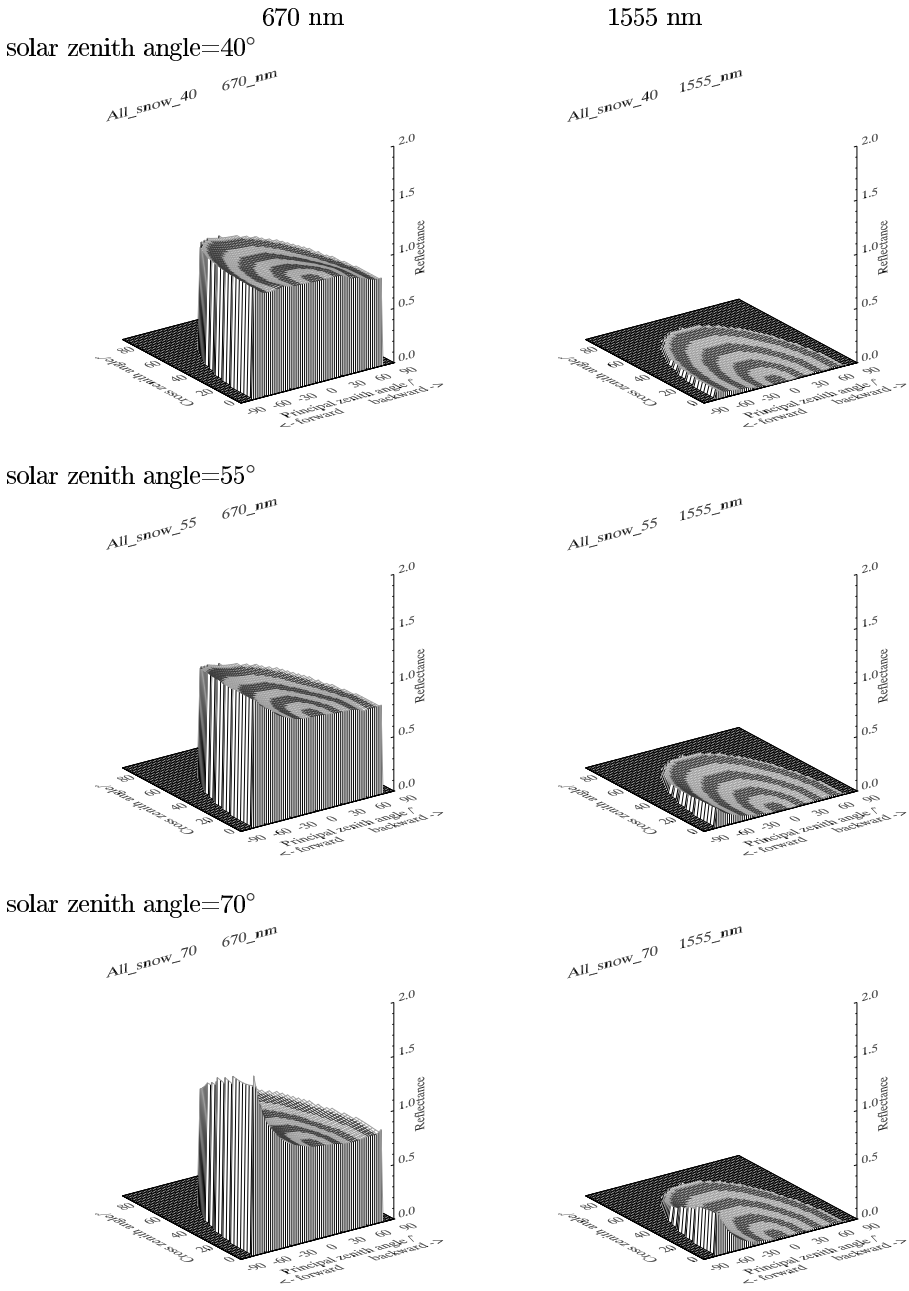


Fig. 9.10. The general 3-D shape of the snow BRF in varying angles of incidence, from top to bottom: 40°, 55°, 70°) and two wavelengths left 670 nm, right 1555 nm. The results are averaged over all measured samples, and show a typical strong forward scattering pattern, stronger in infrared than in visual, and stronger in larger zenith angles than smaller ones. Minor oscillations in the data are mostly caused by the small footprint size vs. surface roughness, and to some extent an averaging artifact. [For a color version of this figure, see the color section.]

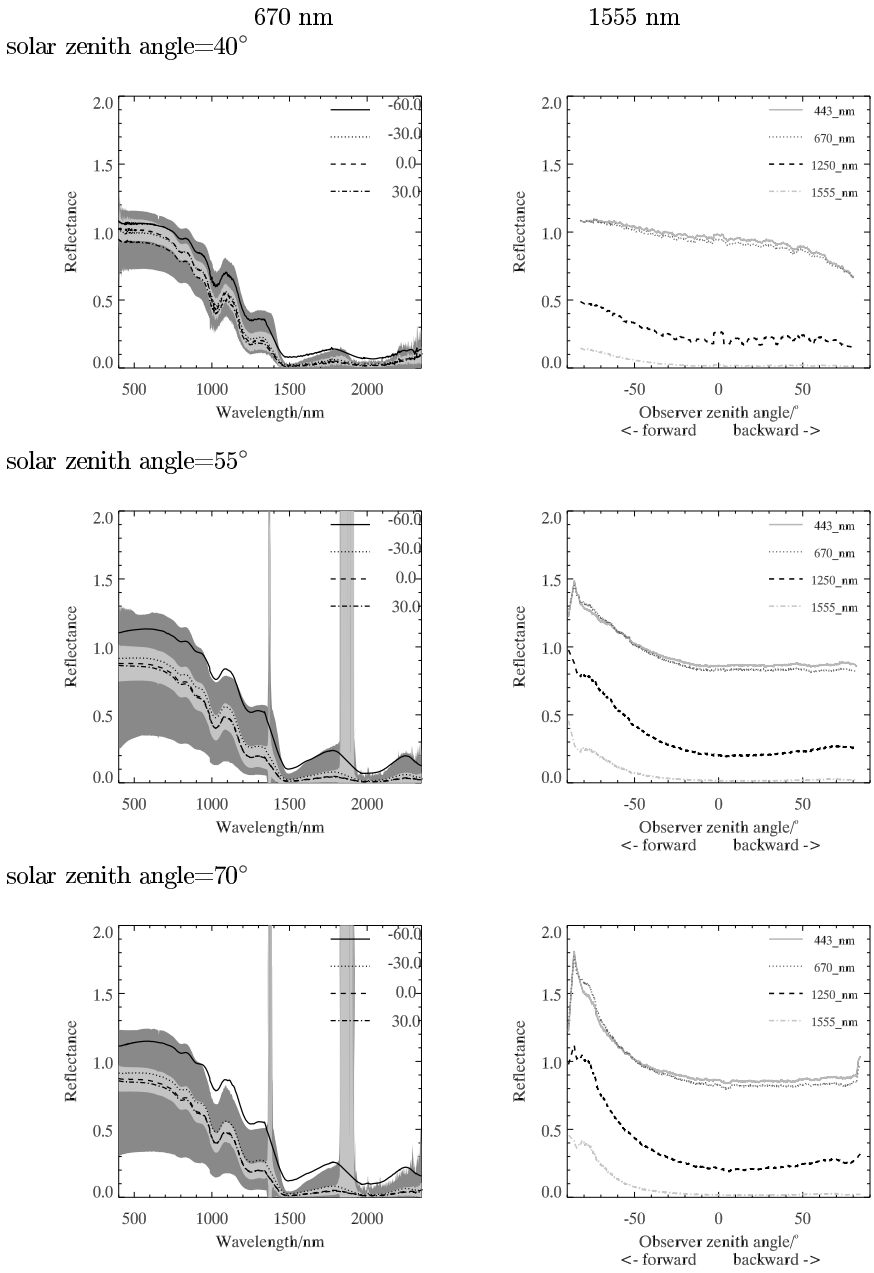


Fig. 9.11. Continuation of Fig. 9.10 with more details. On the left is the reflectivity spectra at four zenith angles on the principal plane (60° forward, 30° forward, nadir, and 30° backward. On the right is the BRF curve on the principal plane as a function of the zenith angle in four channels (violet 445 nm, red 670 nm, SWIR 1225 nm, SWIR 1555 nm). Green colour gives the standard deviation of all measured values at nadir and orange min/max values. The strong peaks and oscillations in some parts of the spectra are sensor noise, because the atmospheric water vapor absorbs all the incident signal in those wavelengths. [For a color version of this figure, see the color section.]

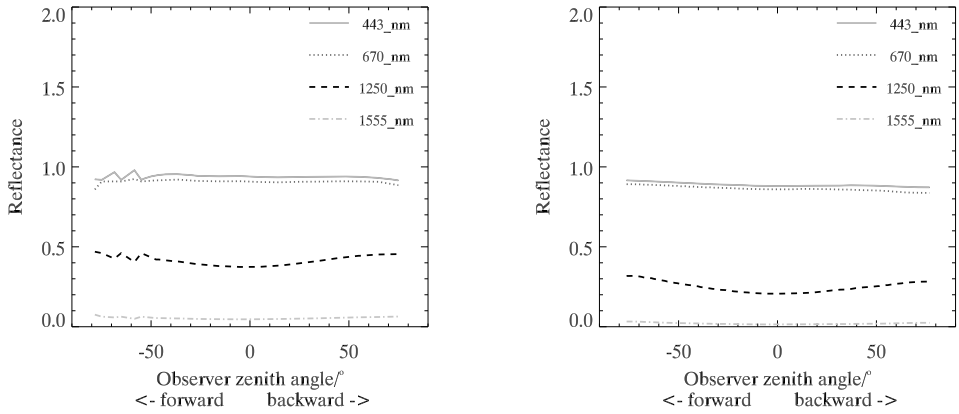


Fig. 9.12. BRF in the perpendicular (cross) plane of two samples, on the left from Pyhätunturi 2005, on the left from Mantovaaranaapa 2009, rough snow sample.

the brightness increases to the furthest forward angle. There is usually no significant backscattering effect except a potential narrow spike less than 1° wide (Kaasalainen et al., 2005). The BRF clearly depends on the angle of incidence. The forward spike increases when the solar zenith angle increases. Azimuthally snow BRF behaves smoothly. All features are best seen in the principal plane. Snow is very white in visual wavelengths, turning very dark in infrared. The spectrum follows the absorption pattern of ice.

Typical polarization is depicted in Figs 9.13 and 9.14. In general, polarization is weak in visual bands, and gets stronger in infrared, where absorption is stronger and single reflection dominates. Polarization is always strongest at forward. In dark wavelengths there appears to be a maximum between 95° to 110° phase angles, but it is not so easy to locate it accurately. The Brewster angle for ice is around 106°. In brighter bands, polarization increases forward to maximum observed phase angles. There is a weak region of negative polarization on the backward side. All representations of polarization – horizontal and vertical separately, their difference (Stokes Q parameter), linear polarization ($-$Q/I), or total polarization ($\sqrt{Q^2+U^2}$) – reveal similar information showing clear spectral and directional dependence, although for various purposes of course one can be more useful than the other. In this measurement setup, also the U-polarization gets significant values, in the first sight contradicting the predictions of the simulations, but the explanation is that the simulations use the scattering plane as a reference, whereas the measurements use the horizontal–vertical plane. The polarization direction appears to be clearly closer to the scattering plane than the horizontal–vertical plane. Some interesting differences between snow types are explained later.

From the samples three basic snow types have been identified: new freshly fallen snow, old dry snow and melting snow. Comparable results are depicted in Figs 9.15 and 9.16. Following these, dry and wet snow are shown in more detail in Figs 9.17 and 9.18 and then as side by side zooms to emphasize the clearest differences in Fig. 9.21.

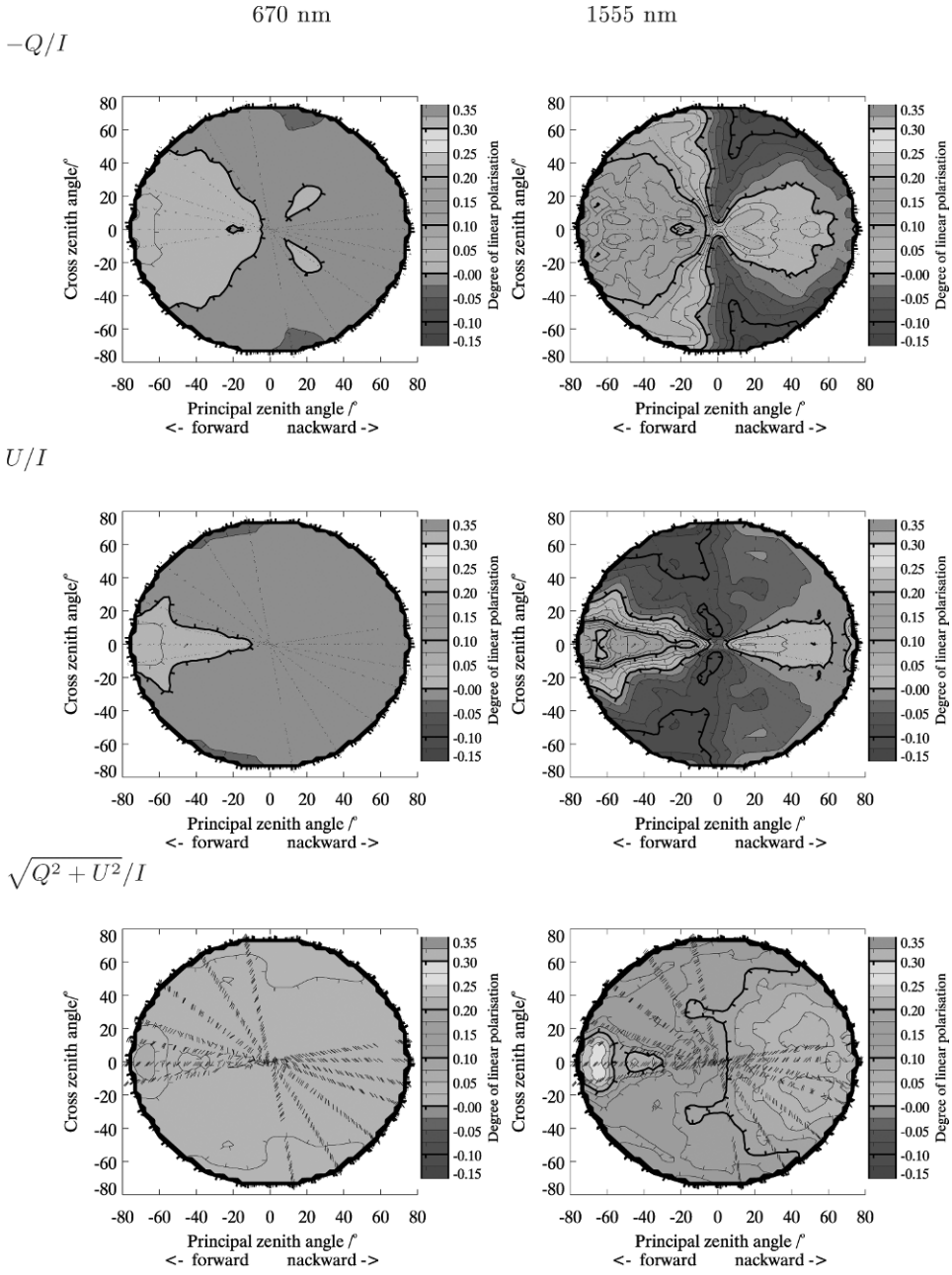


Fig. 9.13. Polarization properties of smooth snow from Mantovaaranaapa 2009. At the top the degree of linear horizontal–vertical polarization (Stokes $-Q/I$), in the middle U/I, and, at the bottom, total linear polarization $\sqrt{Q^2+U^2}/I$, taken at wavelengths of 670 nm (left) and 1555 nm (right), with an angle of incidence of 60°. The short lines in the bottom diagrams indicate the direction of polarization. [For a color version of this figure, see the color section.]

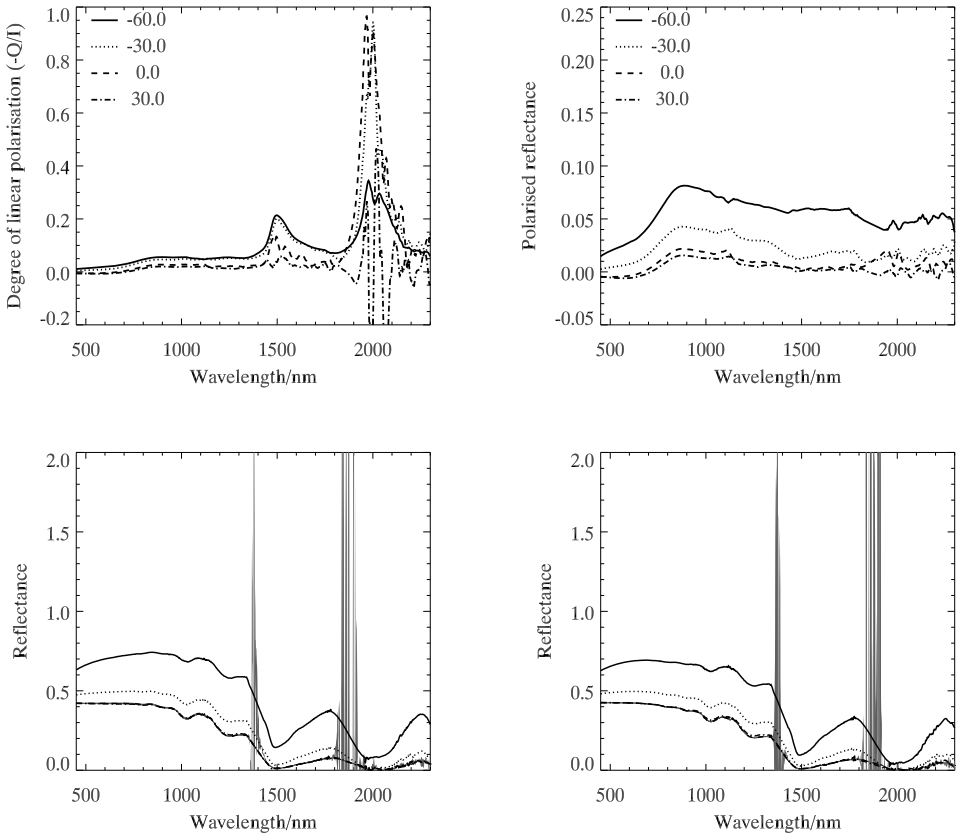


Fig. 9.14. The polarization spectrum shown in different ways. At the top left the degree of linear polarization (−Q/I), at right the polarized reflectance (−Q), and at the bottom horizontally and vertically polarized reflectances separately. All of these show interesting spectral features. [For a color version of this figure, see the color section.]

Below, some features are discussed in more detail, paying attention to how these could reveal some snow properties or identify some snow types (new–old–melting).

9.7.1 Forward scattering signatures

As a forward scattering effect we consider the overall brightening towards larger phase angles/larger zenith angles/smaller scattering angles, not to be confused with the specular effects around the specular angle. All snow types brighten to forward. The forward spike gets stronger when the sun is low, and almost disappears when the sun is high, i.e., it is clearly a phase angle effect. Especially in infrared, snow is a very strong forward scatterer, and in the darkest wavelengths the backward scattering is near the observation threshold.

The most dominant factor for the strength of the forward signature appears to be the roughness of the snow layer: the rougher the surface the smaller the forward scattering. The forward differences in Fig. 9.18 are at least partially due to

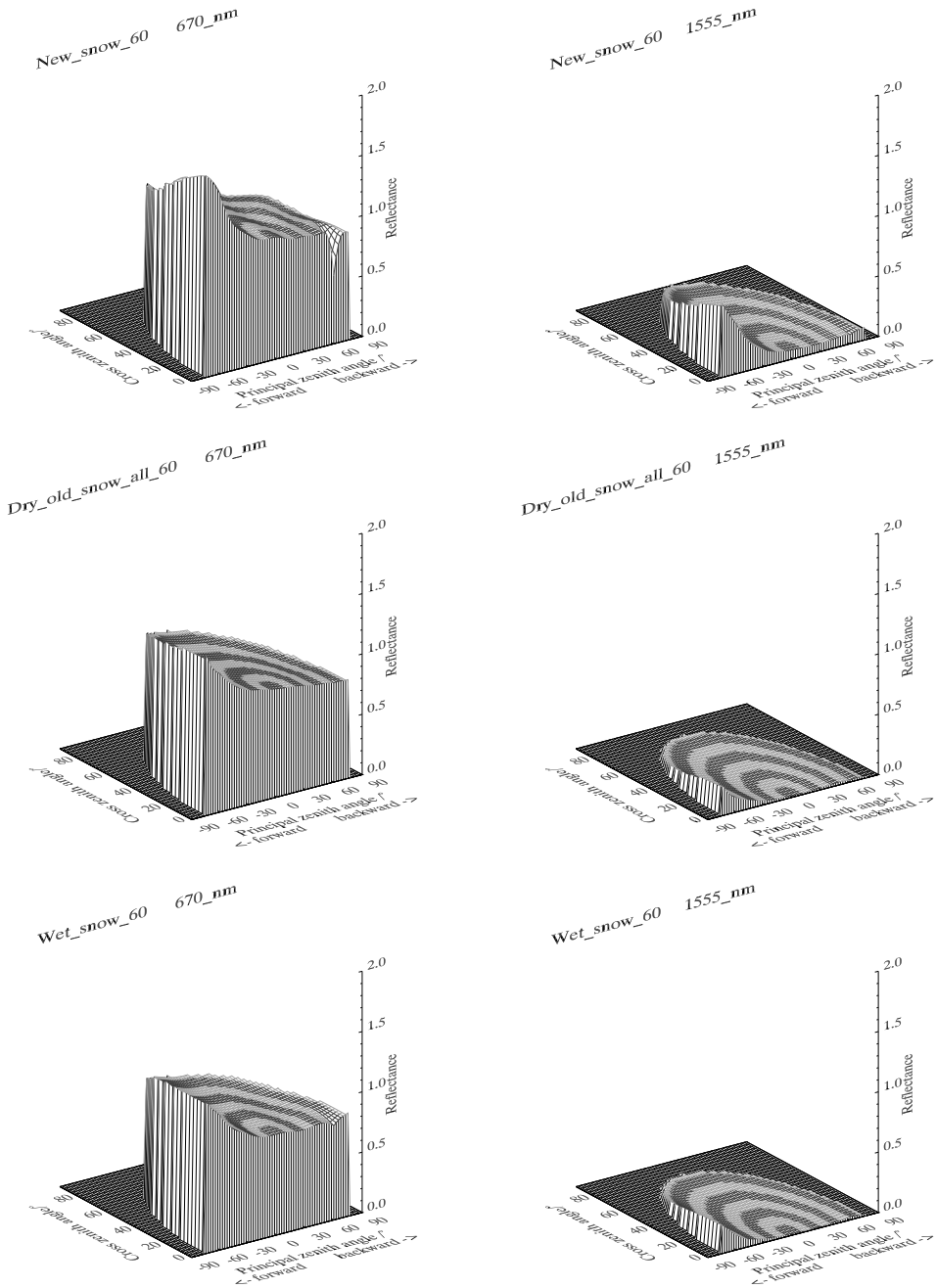


Fig. 9.15. The differences over the main snow types, new snow, old dry snow, and melting snow (from top to bottom), averaged over all samples in the class, otherwise as in Fig. 9.10, with an angle of incidence 60°. [For a color version of this figure, see the color section.]

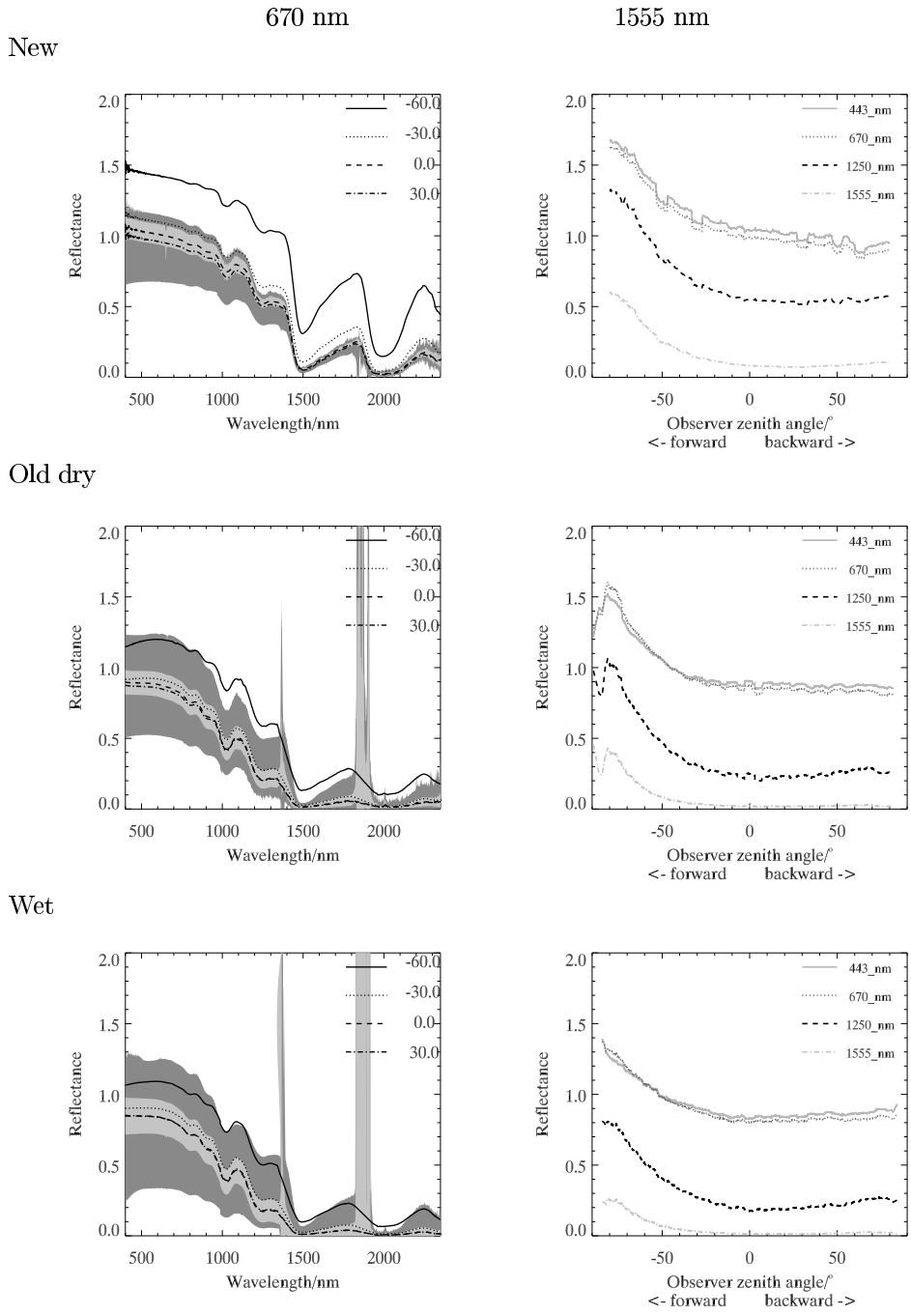


Fig. 9.16. Continuation of Fig. 9.15, analogous to Fig. 9.11. Here one can see more clearly differences in forward scattering and size effect in spectrum. [For a color version of this figure, see the color section.]

2005

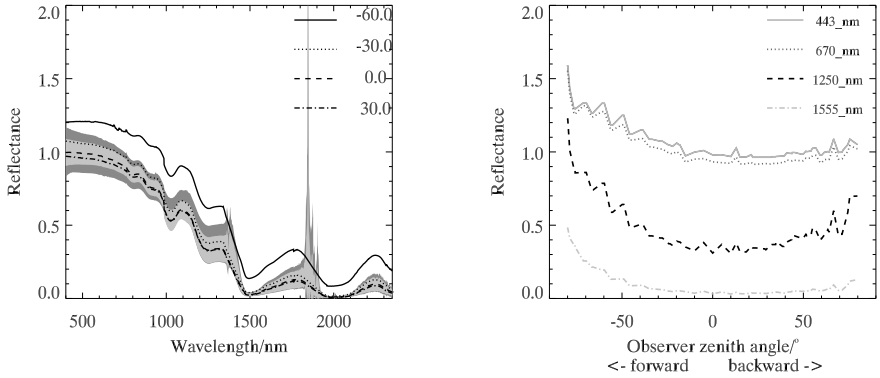
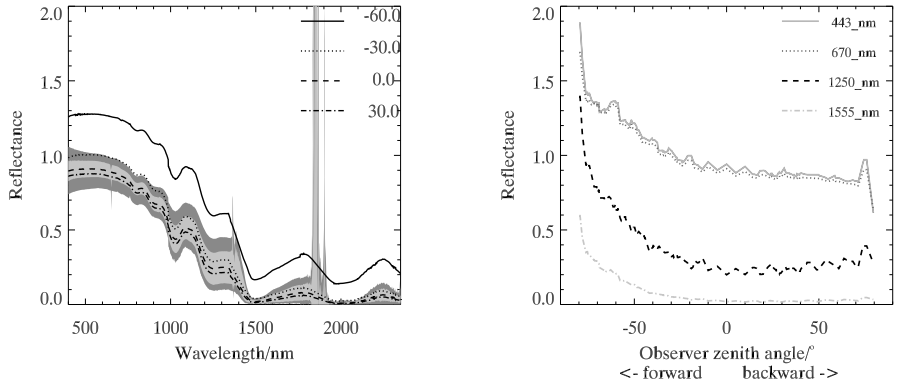
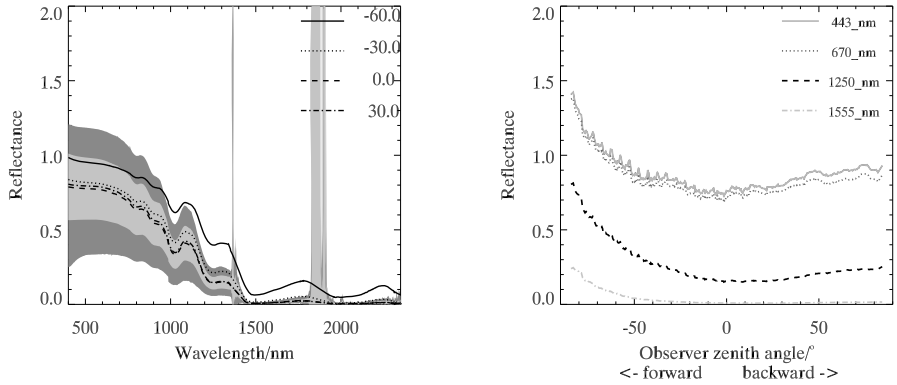


Fig. 9.17. Various dry old snow types, from top to bottom, 2005, 2007, 2008, as in Fig. 9.11, and with an angle of incidence 60°. From this figure one can already see, that samples differ, and single dry snow class is a clear oversimplification, but to keep this review in schedule, further dry snow analysis is dropped. [For a color version of this figure, see the color section.]

2005



2007



2009

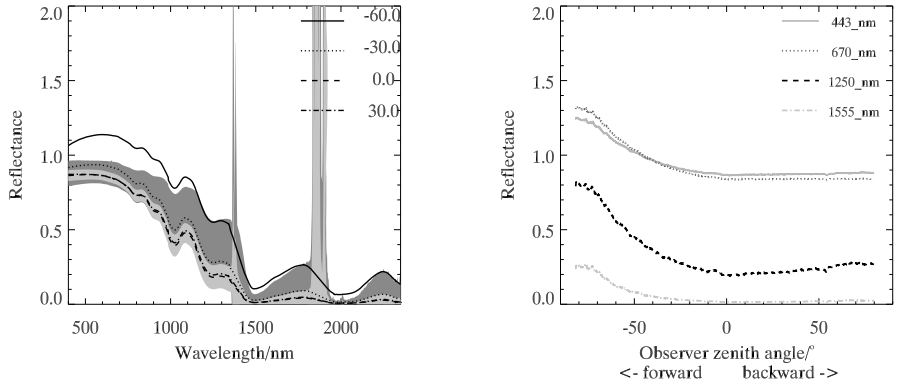


Fig. 9.18. Various wet and melting snow types, from top to bottom 2005, 2007 and 2009 snows, as in Fig. 9.11, and with an angle of incidence 60°. Note especially the small bump and drop at 60° to 70° forward. [For a color version of this figure, see the color section.]

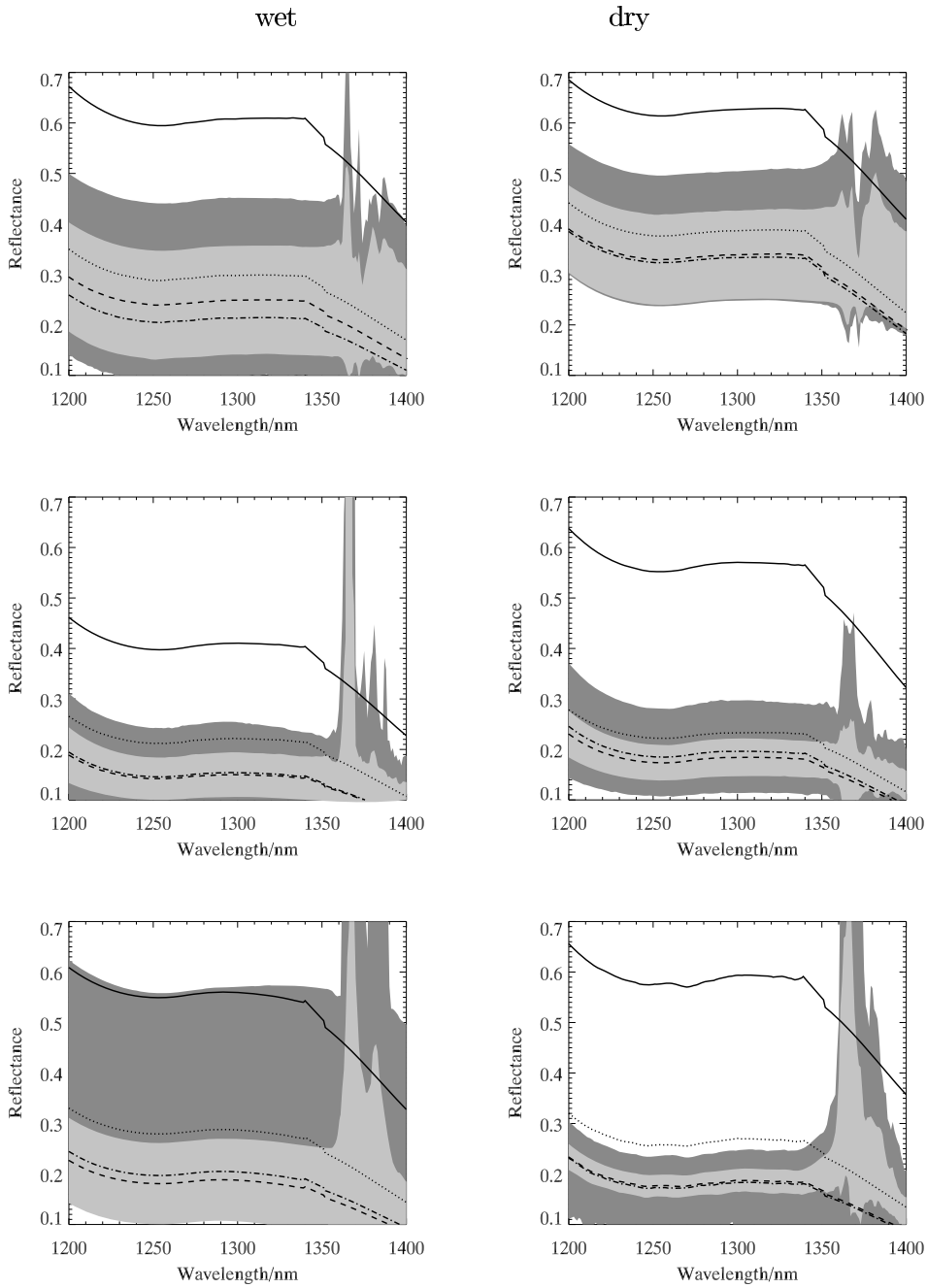


Fig. 9.19. More detailed view of the reflection spectra in the most wetness sensitive region, on the left wet snow, on the right dry old snow. The most indicative feature is the relative difference between 1250 and 1350 nm. [For a color version of this figure, see the color section.]

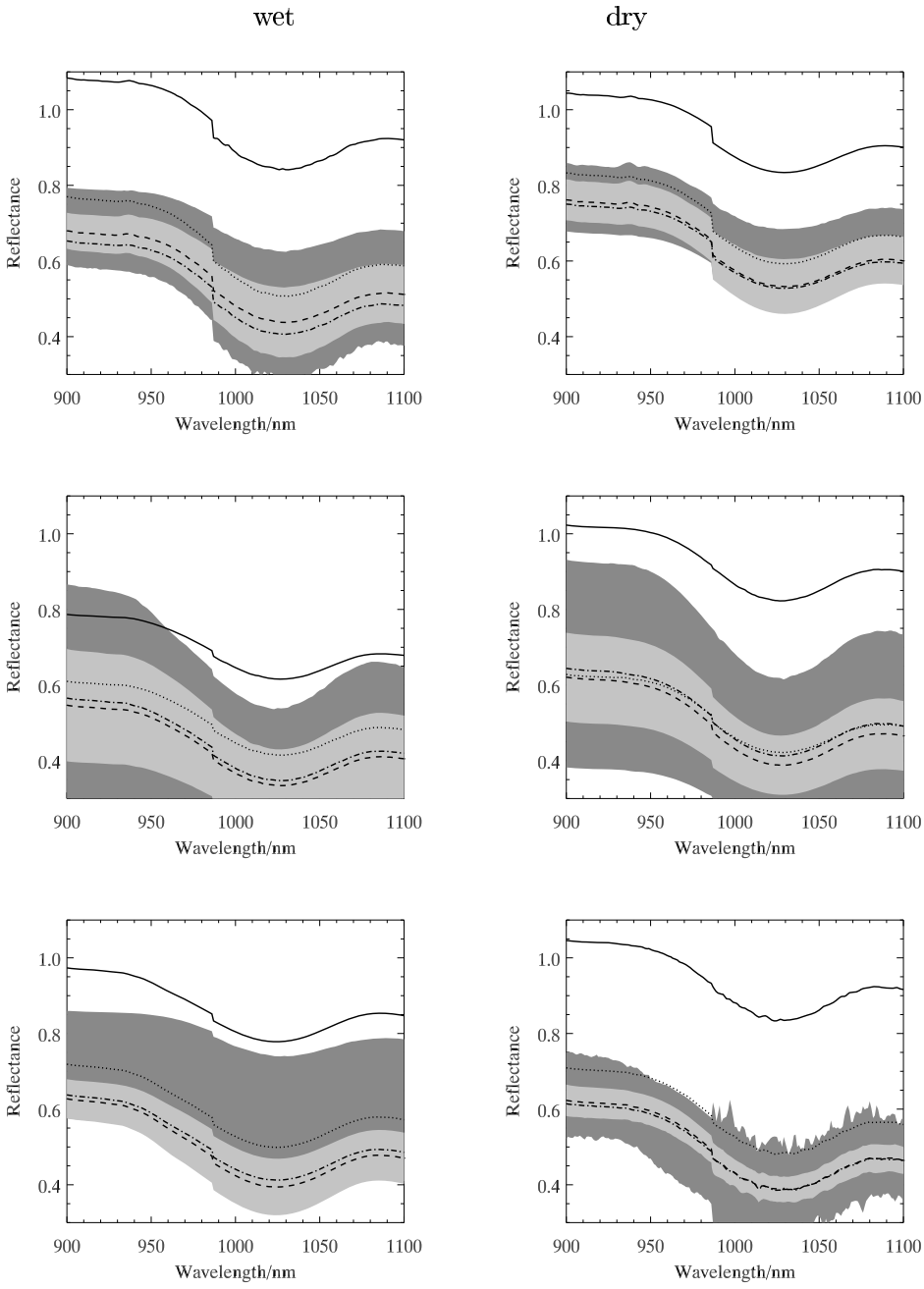


Fig. 9.20. Fig. 9.19 continues at spectral range 900 to 1100 nm. The most indicative feature here should be the difference between 960 nm and 1000 nm or the location of the minimum, although it is very difficult to see even from a zoomed curve, and numerical comparison is needed. Warning that the sensor border is here at 981 nm, which typically causes an erratic step of 1% to 5%, confusing analysis. [For a color version of this figure, see the color section.]

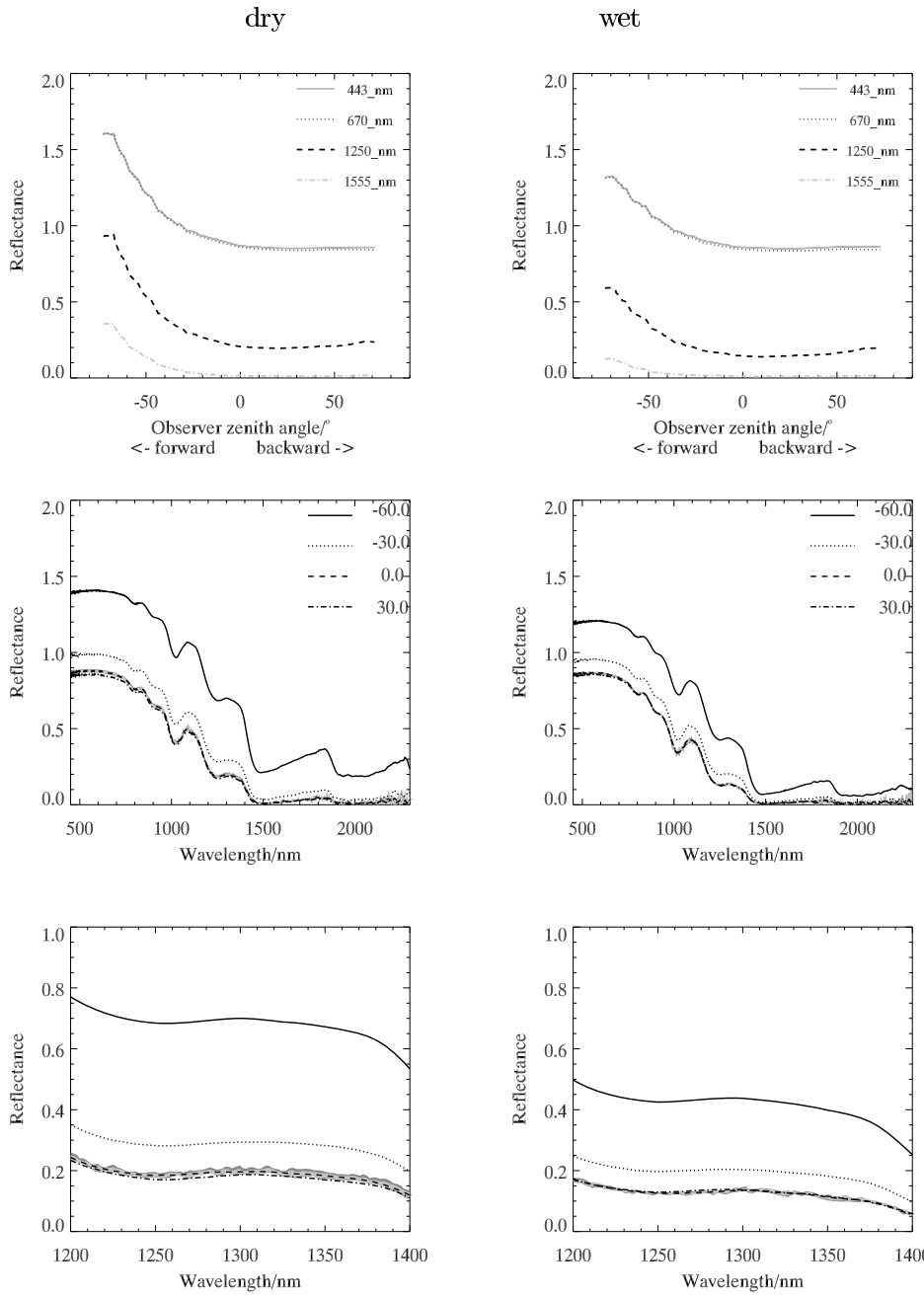


Fig. 9.21. Monitoring the melting process in laboratory. The same snow when dry (left) and after about 1 hour melting inside (right). At the top there is the BRF on the principal plane, in the middle the spectral plot from four directions in the principal plane, and at the bottom a zoom into the more wetness-sensitive spectral band. The light zenith angle was 60°. [For a color version of this figure, see the color section.]

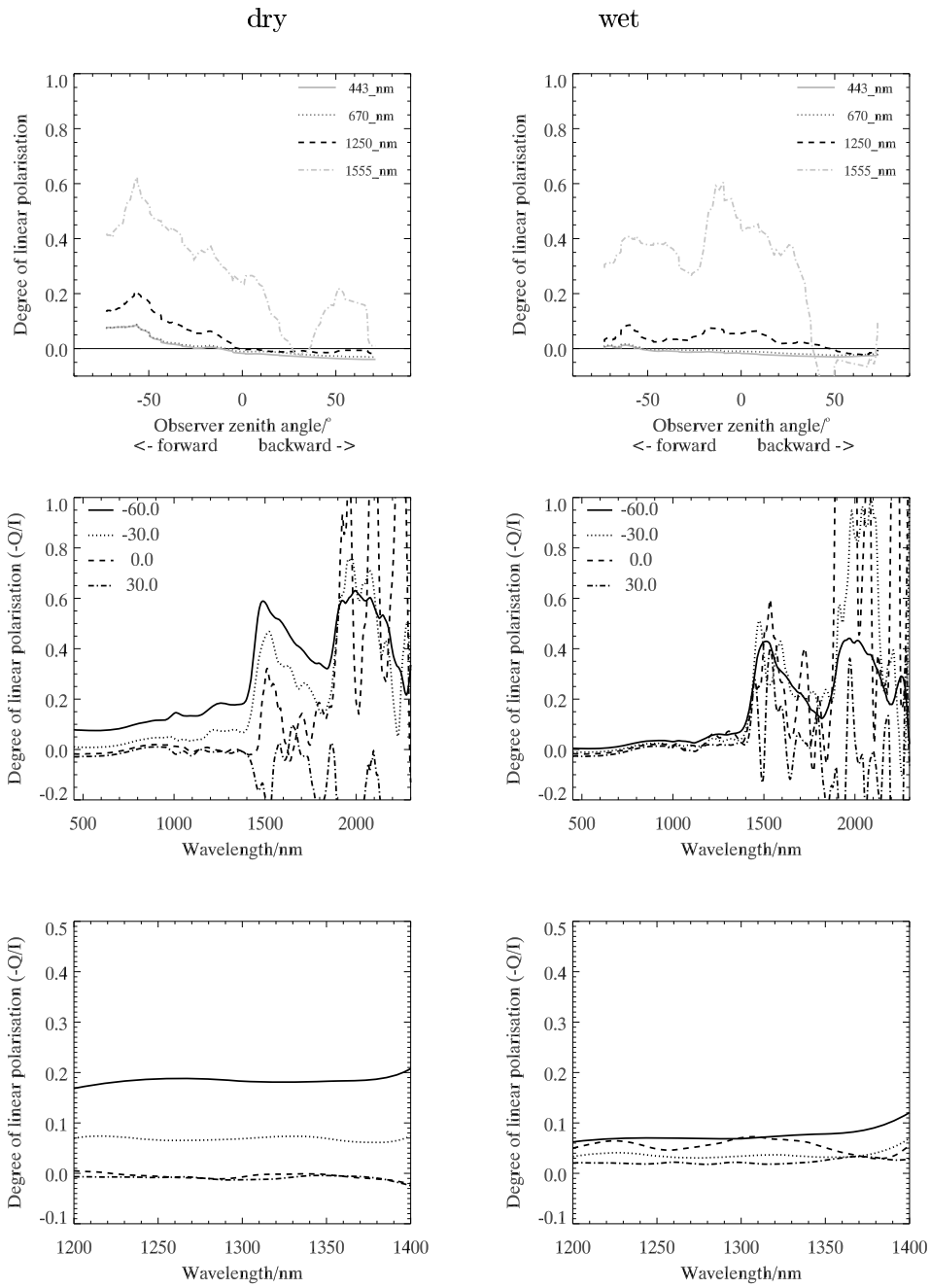


Fig. 9.22. Monitoring the melting process continues. Top: degree of linear polarization on the principal plane (the 1555 nm curve is already very inaccurate outside forward direction); middle: polarization spectrum in four directions (again, from 1400 nm to 2400 nm mostly noise outside forward direction); bottom: zoom to the most wetness sensitive band.

roughness: the 2005 snow was rather flat, whereas the 2007 and 2009 snows were very rough. Warren et al. (1998) also noted a clear roughness dependence.

Wetness also appears to decrease the forward scattering in varying amounts, e.g., Figs. 9.15 and 9.21, but the effect of wetness seems smaller than the effect of roughness, and thus forward scattering alone does not create a unique melting signature. The reduced forward scattering appears to be connected to reduced albedo, i.e., the light that disappears from forward direction is not distributed to other directions, but is absorbed or transmitted to ground. Decrease of snow albedo during warming has been observed in several independent albedo measurements (Pirazzini, 2008, and references therein).

Other factors, e.g., grain shape or packing density, are more difficult to see from the measurements, although simulations suggest a small dependence, 2–10% (Peltoniemi, 2007). Artificial compressing, flattening, or roughening did change BRF observably (Figs 9.23 and 9.24).

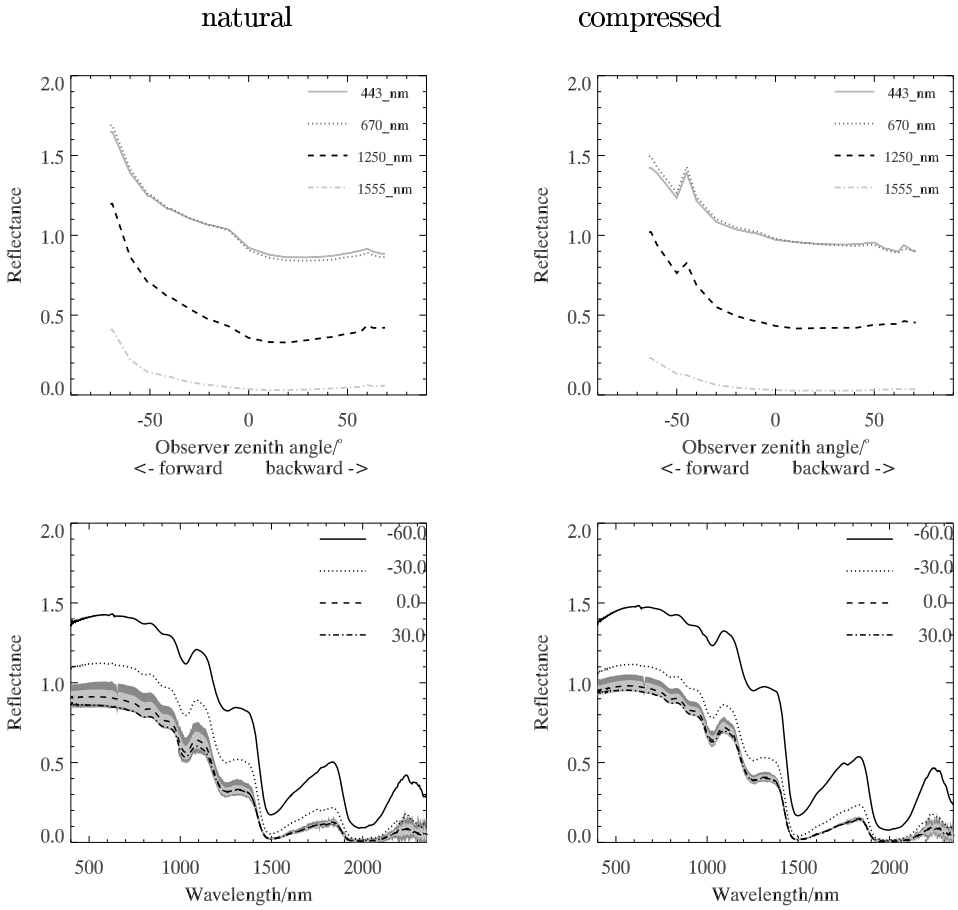


Fig. 9.23. Difference between natural (left) and compressed (right) snow. [For a color version of this figure, see the color section.]

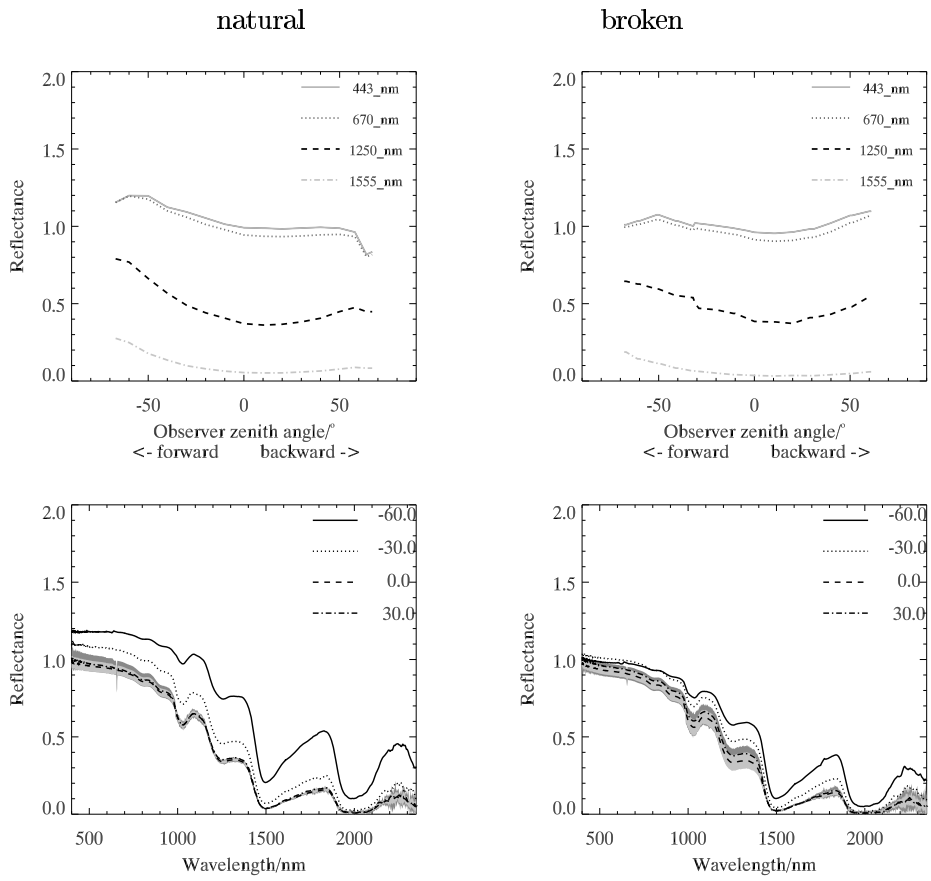


Fig. 9.24. Difference between natural (left) and broken (right) snow. [For a color version of this figure, see the color section.]

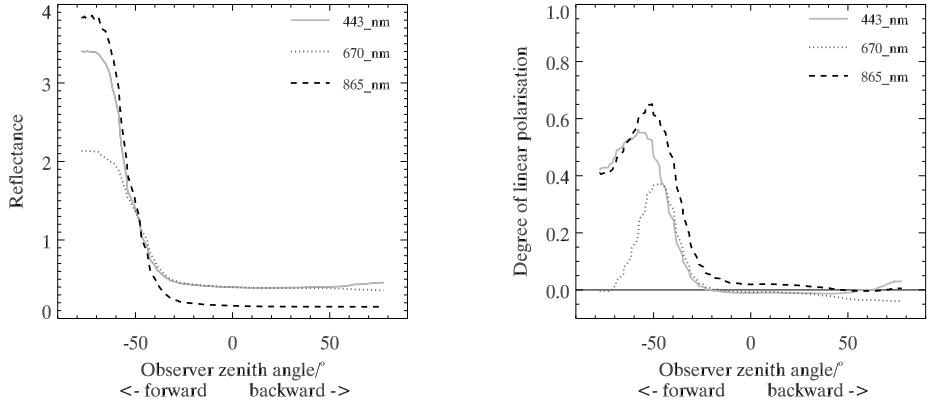


Fig. 9.25. Some measurements of a rough ice layer. Because of sensor saturation in specular and forward direction and very low reflection to other directions, especially in infrared, the results are only partially acceptable.

9.7.2 Specular scattering effects

In some discussions the forward effect is erroneously referred as a specular effect, but pure snow has no specular spike at all. Only when there are liquid or frozen water ponds is a specular spike seen, as shown in Fig. 9.25.

However, in some wet snow data there appears to be some anti-effect, a clear drop a few degrees forward from the specular direction, leaving something that looks like a small bump, or specular spike, but actually the 'bump' just follows the normal forward brightening curve until the drop. The effect is seen well in 1997 (Peltoniemi et al., 2005a), 2005, and 2007 measurements. However, 2009 wet snow samples have only reduced forward scattering without any feature around specular direction. Measuring the same snow frozen in the morning and melting in the afternoon revealed that this directional wetness signal disappears when the snow refreezes again, but cannot be yet confirmed whether the timescale for this is minutes or hours.

The mechanism claiming to explain how the wetness changes the BRF is not proven yet. Simulations cannot yet reproduce this kind of specular feature. There were no macroscopic liquid water ponds in the snow, but there may be some micro-ponds or bridges between snow grains, or there may be some hanging drops under the grains, or there may be a smooth liquid water coating over the grains. Significant metamorphosis during melting is quite to be expected, but reverse metamorphosis during freezing needs more imagination. Visual observation or photography does not give good details, because separating the liquid and frozen phases is too difficult.

9.7.3 Spectral effects

The spectrum is well known to depend on the grain size and all measurements and modeling confirm this. Actually, from the spectral shape one could easily detect even very small, say 1%, differences in a characteristic size, although grain shape effects, sintering and clustering complicate the picture somewhat (what is a grain and its size?). For example, the large spectral difference between new and old snow is mostly a grain-size effect (Fig. 9.16). Further, one can reveal information from the size distributions, e.g., typically one needs some larger grains to increase absorption in 800–1100 nm, and smaller grains to brighten the 1400–2500 nm band.

Another interesting spectral effect is the liquid water signature. Ice and water have refractive indices very close to each other throughout the optical band. The two most indicative bands where the refractive indices of water and ice differ are between 940 and 1030 nm, and between 1250 and 1400 nm. Models can predict the spectral effects easily, and also the field and airborne experiments by Green et al. (2006) have observed it. A zoomed view to the region is shown in Fig. 9.19, but the differences are difficult to see from figures. To quantify the differences, three wetness indices were invented,

$$\text{index1} = 1 - S(1090\,\text{nm})/S(990\,\text{nm}), \tag{9.14}$$

$$\text{index2} = 1 - S(1340\,\text{nm})/S(1255\,\text{nm}), \tag{9.15}$$

$$\text{index3} = 1 - S(1375\,\text{nm})/S(1255\,\text{nm}). \tag{9.16}$$

These are selected from the sides of a minimum in such a way that the intensities are approximately at the same level, to minimize the effects of grain size. Also, because of the sensor border of the ASD spectrometer at 989 nm, the first index is taken a little bit narrower than optimal. The third index is strongest, but overlaps with water vapor absorption, thus it is useful only for lamp measurements. All of them should be small or negative for dry snow, and positive for wet snow. The indices, averaged over a zenith angle cone of 33° over nadir to reduce noise, are shown in Table 9.4. One can see that indeed, most wet and dry snows can be distinguished spectrally, from all these three indices. However, there are some contrary cases too. Dry_old_snow_040316 and Dry_old_snow_080401 look wetter than estimated. It is possible that the lamp heated the snow enough to initiate melting in these cases. Wet_snow_050210 shows mixed effect, which could also easily be explained by the red container. Wet_snow_050419 was initially wet, but was freezing during the night measurement, which might already have influenced the result. Wet_snow_050423 was measured on a warm and sunny day when snow started feeling wet on the surface, but it could be that internals were still dry. The successive measurements at Mantovaaranaapa and SnowFGILab_090320 starting from dry conditions and ending with very wet samples then nicely confirm the power of these indices.

The zenith angle dependence is plotted in Fig. 9.26. There is some dependence, the backward angles being the most sensitive. There is rather a lot of noise in the data, requiring sufficient integration for target and reference measurements. Some discrepancies above could even come from white reference measurement.

It is unfortunate that the ASD sensor border is located exactly in the first wetness sensing band. There is often a 1–10% randomly systematic step in this border, preventing quantitative analysis over this border effectively. Also, the remaining parts of this region are in the less sensitive part of the SWIR 1 detector (Suomalainen et al., 2009a). The second region is in the best parts of the SWIR1 sensor, but here the atmospheric vapor limits the best wavelength range (index 3). For *in situ* measurements, index 3 is no doubt the most sensitive.

9.7.4 Polarization signals

The most distinguishing signal in polarization is that old dry snow polarizes surprisingly strongly in the forward direction, even in visual wavelengths 5% to 35%. New snow or melting snow polarize much more weakly in visual bands, about 0–5%.

The simulation results support very well the observational conclusions, suggesting further that the maximum value of the polarization depends on the grain size, and its location depends very much on the real part of the refractive index, i.e., it could be used to invert for the composition of the scatterer (but in practice this is very demanding). From the simulations one can predict that the 45°/135° linear polarization (in the scattering plane) and the circular polarization (Stokes U and V parameters) are close to zero, and brings no significant extra information.

9.7.5 Albedos

Albedos have been been observed to decrease during melting and in diurnal cycles, see, e.g., Pirazzini (2008) and references therein. In the here reported measurements, the average broadband albedo at solar zenith angle of 60° was 0.87 for new

Table 9.4. Some indicators for wetness. First the sample name, sun zenith angle, estimated or measured wetness [%], and the three invented wetness indices. The larger the index, the wetter the snow should be. The last column gives the maximum polarization in visual wavelengths, when measured. The results are in estimated dryness order, except that the ice and laboratory cases are last.

Sample name	Angle of inc.	Wetness	Index 1	Index 2	Index 3	MaxVisPol
New_Snow_020307	72.0	0	−0.005	−0.020	0.096	
New_Snow_020308	72.0	0	−0.004	−0.027	−0.120	
New_Snow_010307	67.0	0–1	0.005	−0.005	0.073	
New_Snow_080404	55.0	0–2	0.001	−0.013	0.020	0.08
New_Snow_080404	35.0	0–2	0.001	−0.015	0.022	0.04
New_Snow_080405	42.0	0–2	−0.004	−0.011	0.027	0.02
New_Snow_080405	55.0	0–2	−0.005	−0.016	0.027	0.03
New_Snow_080405	65.0	0–2	−0.020	−0.013	0.023	0.02
Dry_old_snow040313a	62.0	0–1	−0.011	−0.033	0.020	
Dry_old_snow040313a	45.0	0–1	−0.013	−0.034	0.022	
Dry_broken_snow040313b	45.0	0–1	−0.020	−0.042	0.021	
Dry_broken_snow040313b	68.0	0–1	−0.007	−0.036	0.026	
Dry_old_snow040314a	65.0	0–1	0.003	−0.030	0.025	
Dry_old_snow040314a	43.0	0–1	0.003	−0.033	0.025	
Dry_compressed_snow040314b	45.0	0–1	−0.003	−0.030	0.017	
Dry_compressed_snow040314b	60.0	0–1	0.006	−0.032	0.022	
Dry_old_snow040316	50.0	0–1	0.075	0.057	0.223	
Dry_old_snow040316	65.0	0–1	0.062	0.027	0.171	
Dry_old_snow050421	73.0	0–1	−0.015	−0.036		
Dry_old_snow050419	60.0	0–2	−0.006	−0.022	−0.000	
Dry_old_snow07	60.0	1–2	0.018	−0.036	0.113	
Dry_old_snow_080401	63.0	0–2	0.075	0.048	0.203	
Dry_old_snow_080401	43.0	0–2	0.046	0.025	0.175	
Dry_old_snow_080402	66.0	0–2	0.022	−0.030	0.056	0.35
Dry_old_snow_080402	45.0	0–2	0.021	−0.049	0.044	0.12
Dry_old_snow_080402	37.0	0–2	0.021	−0.049	0.044	0.05
SnowMasalaField_090319a	66.0	1–2	0.018	−0.015	0.073	0.02
SnowMasalaField_090319b	63.0	1–2	0.014	−0.025	−0.115	0.02
Dry_old_snow050420	60.0	1–3	−0.000	−0.035	−0.006	
Wet_snow050423_60	58.0	1–2	0.002	−0.025	−0.032	
Kommattivaara	58.0	1–3	0.004	−0.019	0.022	0.05
Wet_snow040310	54.0	1–3	0.010	−0.032	0.034	
Wet_snow040311a	72.0	1–3	0.018	0.000		
Wet_snow040311b	62.0	1–3	−0.008	−0.053	0.026	
Wet_snow040311b	44.0	1–3	−0.004	−0.055	0.027	
Mantovaaranaapa_smooth	65.0	1–3	0.007	−0.024	−1.468	0.03
Wet_snow050419	64.0	3–1	0.005	−0.055	0.057	
Wet_snow050419	45.0	3–1	0.004	−0.055	0.059	
Mantovaaranaapa_rough1	60.0	2–4	0.049	0.036	−0.177	0.04
Wet_snow070417_58	58.0	2–5	0.082	0.091	0.434	
Wet_snow070420_60	60.0	2–5	0.035	−0.019		0.04
Korppiaapa_rough	65.0	3–5	0.047	0.016	−0.026	0.03
Mantovaaranaapa_rough2	56.0	3–5	0.068	0.070	0.032	0.03
Wet_snow010504	60.0	2–6	0.074	0.079	0.698	
Wet_snow050210	60.0	3–6	0.022	0.041	0.205	
Wet_snow070418_60	60.0	3–6	0.042	−0.015	0.104	
Mantovaaranaapa_rough3	55.0	4–6	0.080	0.085	0.141	0.02
Wet_snow010505	60.0	4–6	0.078	0.068	0.434	
Wet_snow010508	60.0	4–6	0.079	0.064		
Korppiaapa_Ice	60.0	0	0.175	0.005	−0.644	0.55
SnowFGILab_090320a	60.0	0–2	0.033	0.021	0.169	0.09
SnowFGILab_090320b	60.0	2–4	0.062	0.074	0.263	0.05
SnowFGILab_090320c	60.0	4–6	0.057	0.060	0.233	0.02
SnowFGILab_090320c_72	72.0	4–6	0.069	0.062	0.219	0.05
SnowFGILab_090320c_45	45.0	4–6	0.074	0.044	0.214	0.02

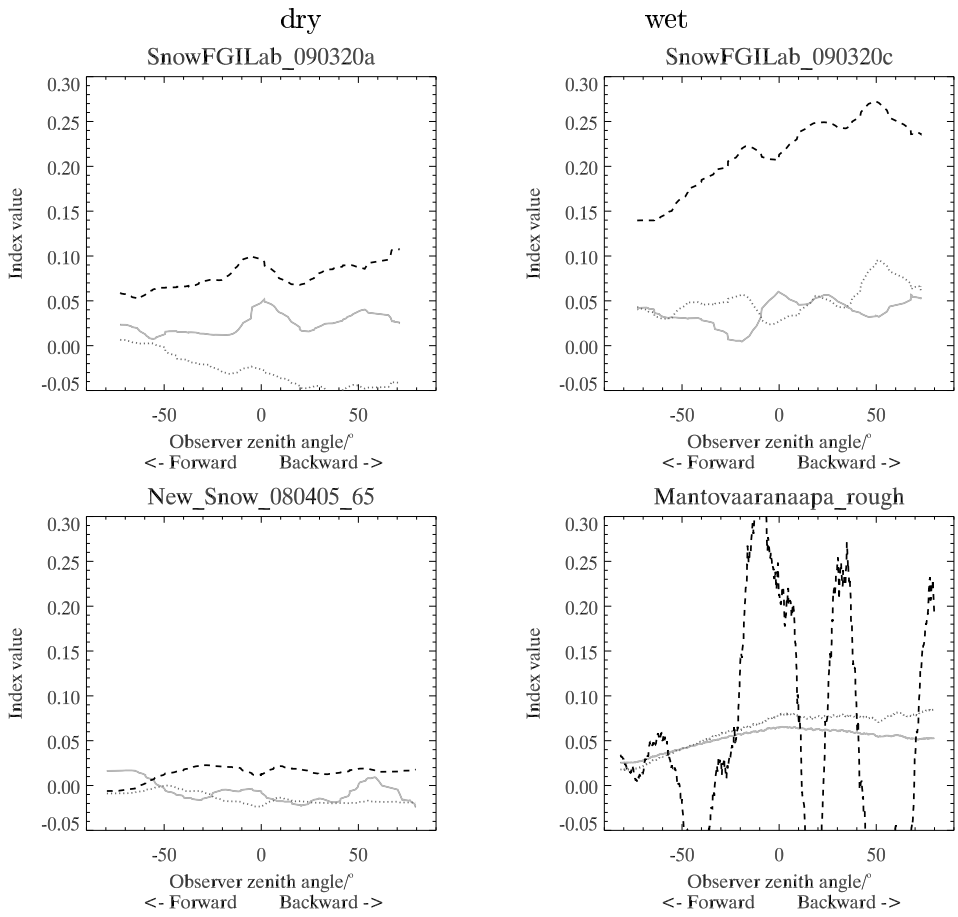


Fig. 9.26. The three wetness indices [1:green, 2:red, 3:black] as a function of the zenith angle when observing four snow samples. Dryer snows are on the left, wetter ones on the right. There is some directional dependence, and most optimum observation direction is from nadir to backwards. In sunlight measurements (Mantovaaranaapa) the index 3 contains mostly only noise, because it overlaps with atmospheric vapor absorption. [For a color version of this figure, see the color section.]

snow, 0.70 for dry old snow, and 0.63 for wet snow. The new and old snow difference is mostly a grain-size effect. Also, the wet snow samples here have probably larger grains than dry snow, so part of the difference could be explained by grain size, and part by the fact that melting snow samples were generally dirtier than dry snows. However, a specific wetness effect was observed even in samples measured wet and dry. Further analysis of albedos, and their dependence on snow properties, angle of incidence, and measurement artifacts is left for forthcoming publications.

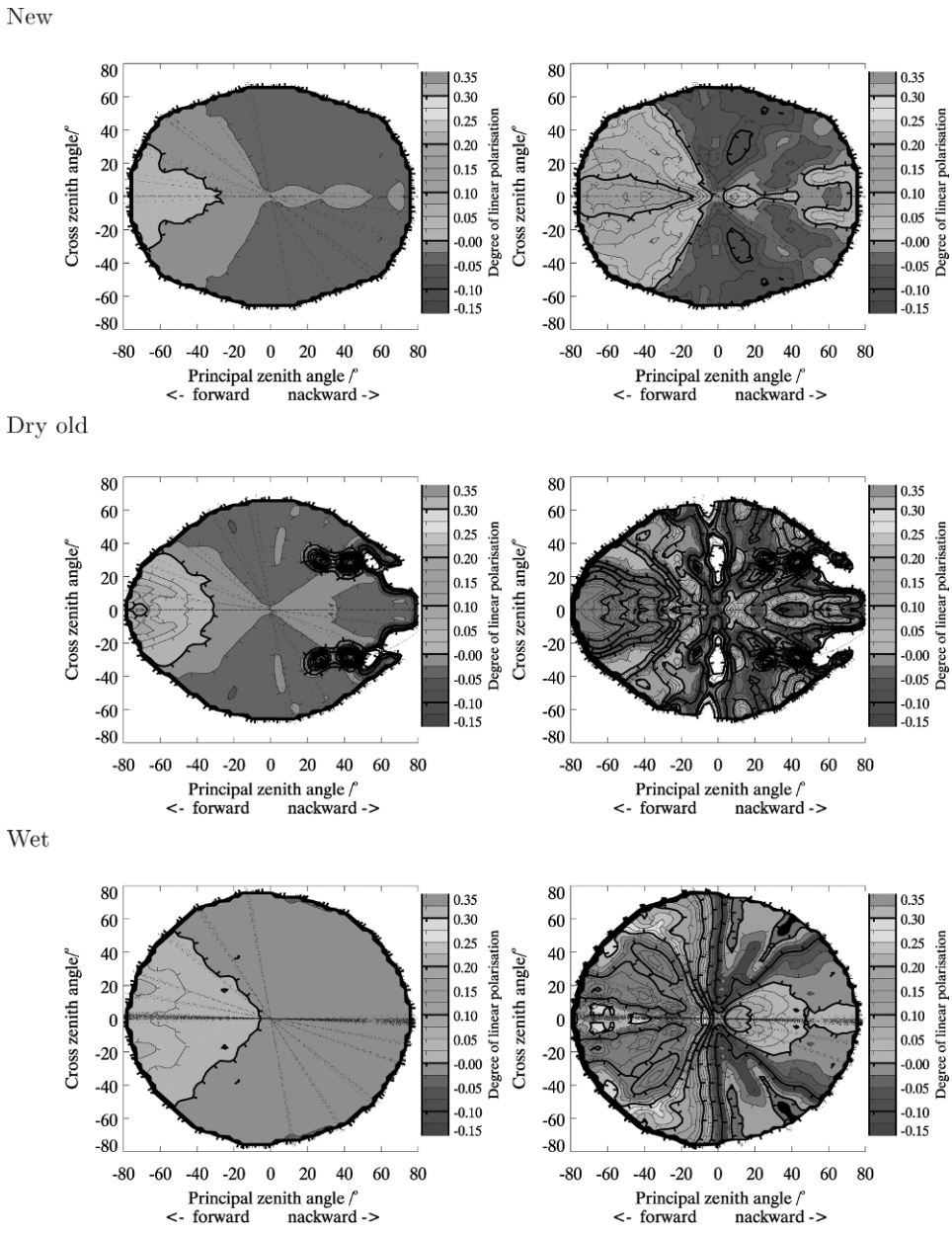


Fig. 9.27. The polarization of three snow types (from top to bottom) new snow, old snow and melting snow. On the left, a contour diagram of the linear polarization ($-Q/I$ only), in 670 nm, on the right in 1555 nm. The angle of incidence was about 65°. [For a color version of this figure, see the color section.]

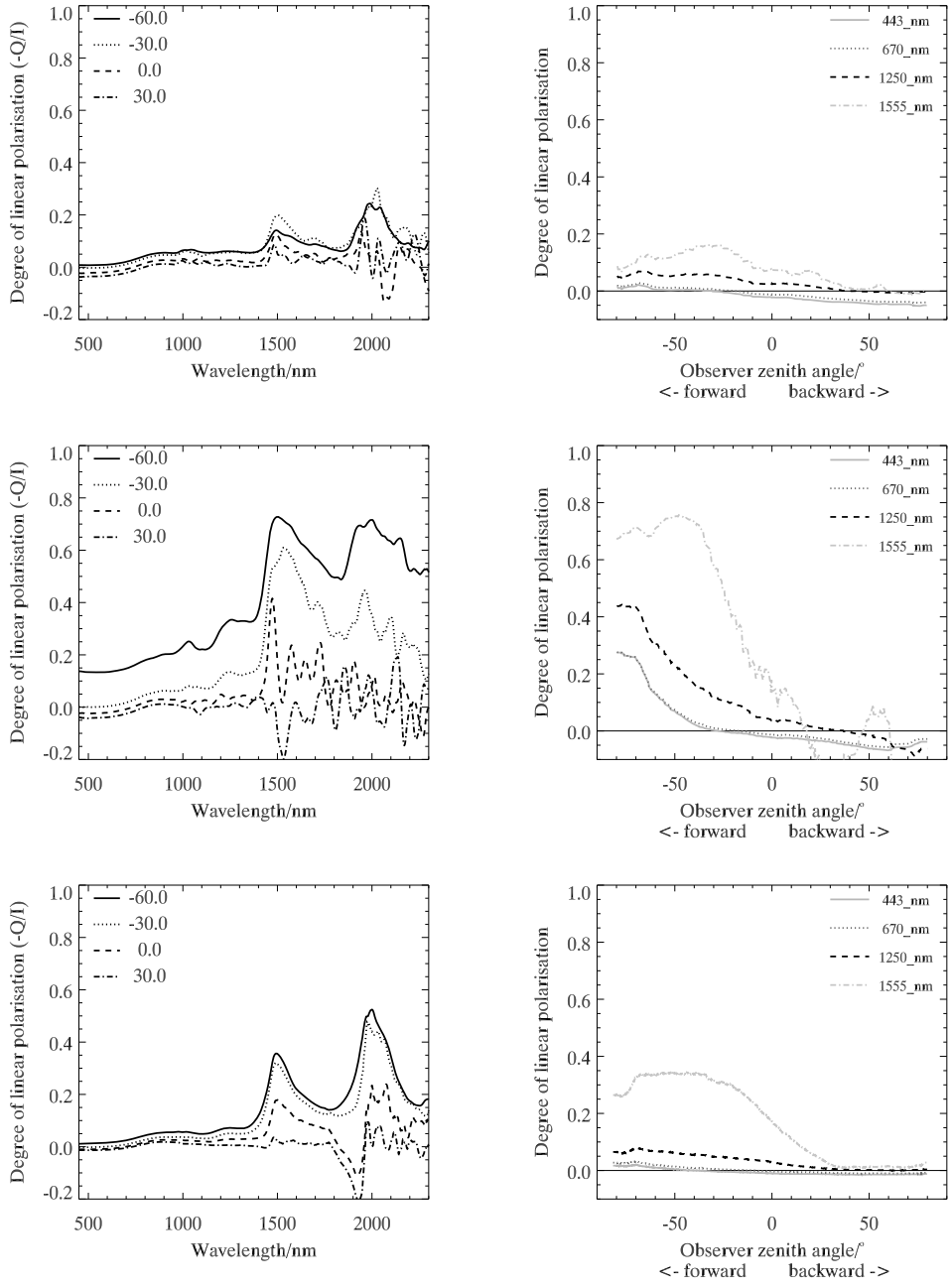


Fig. 9.28. Continuation of Fig. 9.27. On the left the polarization spectra in four zenith angles on the principal plane, and on the right the polarization curves on the principal plane in four wavelengths as a function of the zenith angle.

9.8 Discussion

9.8.1 Melting signatures – a summary

Since monitoring snow melt is one of the most important tasks, the melting signatures are reviewed here. First, measuring the wetness and quantifying the melting of snow is difficult even *in situ* while using the best available techniques. Remote sensing can give a very good idea of the melt line, snow or no snow, but often it is more important to know, when snow starts melting than when it has melted.

The two most promising signals are spectral effects around 1000 and 1300 nm and forward polarization. Both of these can distinguish between completely dry and completely wet snows easily. A more quantitative melting analysis requires still further testing, but there is some promise to get it to an accuracy of 0.01–0.02. The 1000 nm signal has already been validated in real use by Green et al. (2006) and Dozier et al. (2009). Full exploitation of this signal may suffer from the fact that for most typical spectral sensors this range is usually in either end of the sensor wavelength band and thus less sensitive than better bands, or exactly in the border of two sensors, e.g., with ASD FieldSpec at 981 nm. Most optimal sensing thus requires careful selection and tuning of the sensor.

The polarization signal is very promising, but still weakly validated. In the forward direction, melting snow polarizes very little in the visual bands, but old frozen snow polarizes surprisingly strongly, up to 30%. New snow also polarizes weakly, but it can be identified by grain size. How the polarization effect is physically related to melting is unclear. Quantitative use certainly requires better understanding, both theoretically and experimentally.

The directional signal itself is clear and easily measurable. The wetness reduces the strongest forward peak 10–50%, and sometimes causes a specific feature forward from the specular direction. But reduced forward scattering is also caused by increased surface roughness, and thus forward scattering brightness alone cannot reveal wetness.

Albedo decreases during melting. Albedo depends also on grain size, layer structure, thickness, topography, impurities and many other factors. Thus from a single albedo figure alone very little about melting can be said, but if one can monitor daily variations and exclude zenith angle and surface tilt effects, more possibilities for that should exist.

All optical observations see mainly the topmost layer. Polarization is probably caused by the first surface or grain that the sun shines into and the observer sees, and likewise the exploitable directional signals must come from first- or second-order scattering. The 1300 nm spectral signal also cannot see very many scatterings. The 1000 nm signal might be the only one to see a little bit deeper. Since the inner layers can differ significantly from the top layer, even the best observations tell only so much.

Outside the present review, active radar, laser scanning or photogrammetry can be further used to observe snow layer thickness, and radars and radiometry can detect liquid water in snow.

9.8.2 Development of BRF measurement techniques

Goniometer techniques to measure the bidirectional reflectance factor and polarization have been improved during the process. The most important considerations in snow measurements are that the measurements are fast, that the instrument is lightweight and portable enough that it can be transported and setup in various snow types, and that the footprint is large enough. Large spectral range and high resolution have been found very desirable as features, and polarization yields significantly new information. Also fully (or even semi-) automatic measurement of as many parameters as possible helps making results more consistent, comparable and error-free (manual inspection is still needed, but usually human errors corrupt data more than machine errors).

The two most significant limitations of the current measurement setup are still the measurement speed and footprint size. Although, with a good plan and an established routine, a BRF cake can be measured as quickly as in 10 minutes, increasing accuracy, polarization directions, angles of incidence, azimuthal anisotropy, and sampling size can extend the measurements by hours. Often even 10 minutes is too long of time to get the measurements, e.g., if the atmosphere is varying too much, if snow is melting rapidly, or if there are many targets to measure for a campaign reference. Some ways to increase the speed are

- complete two-axis automatisation (10–30%)
- more optimized plan of what angles and polarizations to measure (10–200%)
- dual spectral sensor for simultaneous incident light monitoring (10–30%)
- faster spectrometer (50–500%)
- more advanced lamp setup (50–100%)
- transport, mounting and setup procedures (a few minutes)

A footprint size of about 10 cm is sufficient for new snow, but older snow often has topography and other structures and contaminants in the scale of 10 cm, requiring typically some 50 cm or occasionally an even larger measurement area. Increasing the footprint more necessarily requires some compromises in some other properties. The possibilities for this include

- wider optics: reduces angular resolution (more conical view) or increases backscattering shadow,
- longer arm: increases weight and complexity,
- scanning over larger areas either turning the sensor or moving the whole instrument: often increases measurement time,
- rotating sample trays: ideal and proven for laboratory work, not possible for undisturbed natural targets.

Many of these also solve simultaneously the disturbances of footprint elongations at higher zenith angles and the heterogeneous field of view problems. As is apparent, increasing measurement speed can also be used to improve sampling.

Using an array of sensors, or imaging techniques with light fibers or other optical tricks, could yield snapshot measurements. Alternatively, some optimised combination of a high-resolution spectrometer, a mid-resolution polarimeter, and medium to low spectral resolution cameras might yield almost the same results more affordably and flexibly. For even larger area measurements, a promising approach is

Fig. 9.29. Experimenting with BRF measurements simultaneously using FIGIFIGO (on the right of the background) and Microdrone md4-200 UAV with Ricoh GR II digital camera. The pilot stands on the left of the foreground, and in the middle of the background is another spectral measurement going on. [For a color version of this figure, see the color section.]

small unmanned aerial vehicle systems (UAV) carrying a camera or spectrometer, capable of measuring reflectances of areas of about 1 m^2 to 1 km^2 and many other things.

Also, developing more ideal laboratory and field light source needs effort.

9.8.3 Supporting snow measurements

In order to calibrate remote sensing techniques and systematise BRF measurements, one needs accurate measurement and documentation of all relevant snow properties. Unfortunately, this is much more complicated than appears at the first sight. The complicating factors include rapid metamorphosis of snow, corruption of the sample during measurements, lack of portable precision instruments, rich 3-D variations etc. Even trivial-sounding measurements can be complicated. For example, snow thickness can be easily measured using a simple rod, but if the underlying topography or vegetation varies, snow thickness may also vary, without visible signs. Grain size can be measured using sieves, or photographing an ensemble of loose grains. But if the snow is wet, grains stick to each other, and will not go through a sieve and cannot be separated for individual photography. Photographic plate also immediately melts wet grains. And when the wet snow is frozen to a hard layer, there may not even be separate grains any more but a complicated 3-D structure. Snow wetness can be measured at some accuracy using the electronic snow fork (e.g., by Toikka engineering). The system is, however, not accurate enough for the most important surface layer, where the density and wetness profile may change in cm scale or even smaller.

Thus, for the validation of remote sensing techniques, and to support modeling, new independent real time field measurement techniques are needed at least for

- wetness and density profiles in a resolution of 1 cm or better,
- grain shapes and size distributions, or 3-D structures of complex snow layers, also for wet snow or hard snow layers,
- wavelength-scale surface and internal structures of snow grains, e.g., an economic portable winter-proof electron microscope,
- detailed locations of melt water in snow layers (inside, outside, under or between grains?),
- snow and ground surface topography in mm to hm scales,
- snow contaminants and impurities.

Also, setups where one can, in some controlled and reproducible way, change snow properties and systematically measure the effects of various parameters would be extremely useful.

9.8.4 Modeling

The ray optics model can predict well the spectral effects of grain size, and to some extent the wetness effect, but here reality is more complicated. The model also predicts well the general shape of the BRF, but some details need more polishing, e.g., the wetness features in forward. Modeling the topographic effects needs more work.

But the biggest surprise is polarization. Although the general spectral and forward-biased polarization shapes are predicted, the model does not give any explanation for the difference between wet and dry snow. It is certainly more complicated than a small refractive index difference. A natural point of speculation is wavelength-scale structures in the interiors and surfaces of snow grains, assuming that melting smooth the surface, and freezing roughens it. Also, the negative polarization in backward is not reproduced by model.

The challenges for modeling include

- realistic grain shapes, new flakes, rounded grains, cups, needles, hoar,
- realistic size distribution,
- realistic 3-D structure,
- representation of liquid water, over, under or between grains? Evenly distributed, or located on certain local hot spots of focused sunlight?
- including both radiative transfer and electromagnetic effects, large-scale structures and wavelength-scale ripple,
- combining snow metamorphosis models to reflection models.

9.9 Conclusions

The reflectance of snow has been the subject of our review here. Despite its apparent simplicity, snow as a subject of study is extremely rich and challenging. It shows large variation of physical, chemical, and optical properties, in all spatial and temporal scales. Interpreting data is very demanding, especially, when snow is mixed with sticking out ground objects, in partially bare areas, among vegetation, or when covered by forest.

From the reflection spectrum of snow it is possible to extract rather accurate measures of grain size and indications of wetness. The wetness signal is rather faint, but with good instrumentation and calibration is exploitable. Most indicative bands are near 1000 nm and 1300 nm, and observations should be taken from nadir or backwards. The directional signal, BRF, carries information about grain shape. More regular crystals yield stronger forward peak. Rougher grains give more backscattering. Topographic roughness reduces forward scattering and enhances backward scattering. Wetness reduces forward scattering, until melt ponds and flat frozen ice start producing specular reflections. Unfortunately these all mix together and are not separable without further information. Another clear wetness signal seems to be forward polarization. Dry old snow polarizes up to 10–30% even in visual, whereas wet snow polarizes <5%. New also snow polarizes little, but it can be easily identified by small grain size.

Validating and calibrating measurements, models and inversion procedures requires improved tools to accurately and independently measure in the field the key snow parameters: wetness, grain size, and grain shape. New reflectance measurements are still needed from more natural snow, with complex surface topography, internal 3-D structure, contaminants and impurities. Modeling must try to better explain polarization, take into account macroscopic and microscopic structures, and understand the forms of liquid water in snow.

We conclude that, based on these reflectance measurements, improved tools and techniques for snow monitoring can be developed, with the help of spectrometry, polarimetry, and directional observations.

Acknowledgments

Juha Suomalainen and Teemu Hakala are supported by the Academy of Finland. The EU commission Interreg III program has supported this research via the Norsen project 2004–7. The research projects of 2008 and 2009 are connected to the Snortex campaign, thanks to Terhikki Manninen and Jean-Louis Roujean. Tähtelä Research station in Sodankylä is thanked for their hospitality during our research efforts and expeditions. Leena Matikainen assisted in the 2001 snow measurements.

References

Aoki, T., Aoki, T., Fukabori, M., Hachikubo, A., Tachibana, Y., and Nishio, F. (2000). Effects of snow physical parameters on spectral albedo and bidirectional reflectance of snow. *J. Geophys. Res.*, **105**(D8):10219–10236.

Bohren, C. F. and Barkstrom, B. R. (1974). Theory of optical properties of snow. *J. Geophys. Res.*, **79**:4527–4535.

Bonnefoy, N. (2001). *Développement d'un spectrophoto-goniomètre pour l'etude de la réflectance bidirectionelle de surfaces géophysiques. Application au soufre et perspectives pour le satellite Io*. PhD thesis, Université Joseph Fourier, Grenoble.

Bourgeois, C. S., Calanca, P., and Ohmura, A. (2006a). A field study of the hemispherical directional reflectance factor and spectral albedo of dry snow. *J. Geophys. Res.*, **111**(D20108).

Bourgeois, C. S., Ohmura, A., Schroff, K., Frei, H.-J., and Calanca, P. (2006b). IAC ETH goniospectrometer: A tool for hyperspectral HDRF measurements. *J Atmos Ocean Technol.*, **23**(4):573.

Bruegge, C. J., Helmlinger, M. C., Conel, J. E., Gaitley, B. J., and Abdou, W. A. (2000). PARABOLA III: A sphere-scanning radiometer for field determination of surface anisotropic reflectance functions. *Remote Sensing Reviews*, **19**(1–4):75–94. The Second International Workshop on Multiangular Measurements and Models.

Colbeck, S., Akitaya, E., Armstrong, R., Gubler, H., Lafeuille, J., Lied, K., McClung, D., and Morris, E. (1985). The international classification for seasonal snow on the ground. Technical report, The International Commission on Snow and Ice of the International Association of Scientific Hydrology, and International Glaciological Society. Working Group on Snow Classification.

Courréges-Lacoste, G. B., Schaarsberg, J. G., Sprik, R., and Delwart, S. (2003). Modeling of Spectralon diffusers for radiometric calibration in remote sensing. *Optical Engineering*, **41**(12):3600–3607.

Dingman, S. (2002). *Physical Hydrology*. Prentice Hall.

Dozier, J., Davis, R. E., and Nolin, A. W. (1989). Reflectance and transmittance of snow at high spectral resolution. In *Proc. of the International Geoscience and Remote Sensing Symposium '89*, number No. 89 CH2768-0 in IEEE, pages 662–664, Vancouver, Canada.

Dozier, J., Green, R. O., Nolin, A. W., and Painter, T. H. (2009). Interpretation of snow properties from imaging spectrometry. *Remote Sensing of Environment*, **113**:S25–S37. Supplement 1.

Eskelinen, M., Metsämäki, S., Pulliainen, J., Hallikainen, M., and Praks, J. (2003). The use of airborne optical spectrometer data in snow cover monitoring. In *Geoscience and Remote Sensing Symposium, IGARSS, Proceedings*, volume 4, pages 2823–2825.

Flanner, M. G. and Zender, C. S. (2006). Linking snowpack microphysics and albedo evolution. *J. Geophys. Res.*, **111**(D12208).

Georgiev, G. T. and Butler, J. J. (2004). The effect of incident light polarization on spectralon BRDF measurements. In *Proc. SPIE*, **5570**:492–502.

Giardino, C. and Brivio, P. A. (2003). The application of a dedicated device to acquire bidirectional reflectance factors over natural surfaces. *Int. J. Remote Sensing*, **24**(14):2989–2995.

Gray, D. M. and Male, D. H., editors (1981). *The Handbook of Snow*. Pergamon, Toronto.

Green, R. O. and Dozier, J. (1995). Measurement of the spectral absorption of liquid water in melting snow with an imaging spectrometer. Technical Report 95-1, JPL, Pasadena, California.

Green, R. O. and Dozier, J. (1996). Retrieval of surface snow grain size and melt water from AVIRIS spectra. In *6th AVIRIS Airborne Geoscience Workshop Proceedings*, number 96-0653, Pasadena, California.

Green, R. O., Dozier, J., Roberts, D. A., and Painter, T. H. (2002). Spectral snow reflectance models for grain size and liquid water fraction in melting snow for the solar reflected spectrum. *Annals of Glaciology*, **34**:71–73.

Green, R. O., Painter, T. H., Roberts, D. A., and Dozier, J. (2006). Measuring the expressed abundance of the three phases of water with an imaging spctrometer over melting snow. *Water Resources Research*, **42**:W10402.

Grenfell, T. C. and Warren, S. G. (1999). Representation of a nonspherical ice particle by a collection of independent spheres for scattering and absorption of radiation. *J. Geophys. Res.*, **104**:31697–31709.

Grenfell, T. C., Neshyba, S. P., and Warren, S. G. (2005). Representation of a nonspherical ice particle by a collection of independent spheres for scattering and absorption of radiation 3: Hollow columns and plates. *J. Geophys. Res.*, **110**(D17203).

Haner, D., McGuckin, B., and Bruegge, C. (1999). Polarization characteristics of spectralon illuminated by coherent light. *Appl. Opt.*, **38**:6350–6356.

Hapke, B. (1993). *Theory of Reflectance and Emittance Spectroscopy*. Cambridge University Press.

Hendriks, E., Greuell, J., Oerlemans, J., Knap, W., and Stammes, P. (2004a). Anisotropy of reflected solar short wave radiation on a snow surface: Ground measurements and modeling. In *Geophysical Research Abstracts*, volume 6. SRef-ID: 1607-7962/gra/EGU04-A-03426.

Hendriks, E., Greuell, W., Oerlemans, J., Knap, W., and Stammes, P. (2004b). Anisotropy of reflected solar short wave radiation on a snow surface: Ground measurements and modeling. In *Eos. Trans. AGU*, volume 85. Abstract #C33A-04.

Hudson, S. R., Warren, S. G., Brandt, R. E., Grenfell, T. C., and Six, D. (2006). Spectral bidirectional reflectance of Antarctic snow: Measurements and parameterization. *J. Geophys. Res.*, **111**(D18106).

Jin, Z. and Simpson, J. J. (1999). Bidirectional anisotropic reflectance of snow and sea ice in AVHRR channel 1 and 2 spectral regions – part I: Theoretical analysis. *IEEE Trans. Geosci. Remote Sensing*, **37**(1):543–554.

Kaasalainen, S. (2002). *Backscattering of light from solar system ices and regoliths*. PhD thesis, University of Helsinki.

Kaasalainen, S., Peltoniemi, J., Näränen, J., Suomalainen, J., Kaasalainen, M., and Stenman, F. (2005). Small-angle goniometry for backscattering measurements in the broadband spectrum. *Applied Optics*, **44**(8):1485–1490.

Kaempfer, T. U., Hopkins, M. A., and Perovich, D. K. (2007). A three-dimensional microstructure-based photon-tracking model of radiative transfer in snow. *J. Geophys. Res.*, **112**(D24113).

Knap, W. and Oerlemans, J. (1996). The surface albedo of the greenland ice sheet: Satellite-derived and in situ measurements in the søndre strøfjord area during the 1991 melt season. *Journal of Glaciology*, **42**(141):364–374.

Kneubühler, M., Itten, K. I., Rey, S., Schopfer, J. T., and Schneebeli, M. (2008). Towards combined spectrodirectional reflectance measurements and detailed snow microstructure studies. In *5th EARSeL Workshop on Remote Sensing of Land Ice and Snow*.

Knight, C. A. (1979). Observations of the morphology of melting snow. *Journal of the Atmospheric Sciences*, **36**:1123–1130.

Kokhanovsky, A. A. and Zege, E. P. (2004). Scattering optics of snow. *Appl. Opt.*, **43**:1589–1602.

Kokhanovsky, A. A., Aoki, T., Hachikubo, A., Hori, M., and Zege, E. P. (2005). Reflective properties of natural snow: approximate asymptotic theory versus in situ measurements. *IEEE Transactions on Geoscience and Remote Sensing*, **43**:1529–1535.

Kuhn, M. (1985). Bidirectional reflectance of polar and alpine snow surface. *Annals of Glaciology*, **6**:164–167.

Leroux, C., Deuze, J.-L., Coulomb, P., Sergent, C., and Fily, M. (1998). Ground measurements of the polarized bidirectional reflectance of snow in the near-infrared spectral domain: Comparison with model results. *J. Geophys. Res.*, **103**:18721–18731.

Li, S. and Zhou, X. (2004). Modelling and measuring the spectral bidirectional reflectance factor of snow-covered sea ice: an intercomparison study. *Hydrological Processes*, **18**:3559–3581.

Liang, S. (2004). *Quantitative Remote Sensing of Land Surfaces*. Wiley Series in Remote Sensing. Wiley, Hoboken, NJ.

Libbrecht, K. (2005). The physics of snow crystals. *Reports on Progress in Physics*, **68**:855–895.

Lyapustin, A., Tedesco, M., Wang, Y., Aoki, T., Hori, M., and Kokhanovsky, A. (2009). Retrieval of snow grain size over Greenland using MODIS. *Remote Sensing of Environment*, **113**:1976–1987.

Manninen, T. and Stenberg, P. (2008). Simulation of the effect of snow covered forest floor on the total forest albedo. *Agricultural and Forest Meteorology*, **149**:303–319.

Manninen, T., Siljamo, N., and Poutiainen, J. (2006). Validation of the surface albedo product (SAL) of CM SAF in winter conditions. In *Proceedings of the EUMETSAT Meteorological Satellite Conference*, pages 48–54.

Matikainen, L., Kuittinen, R., and Vepsäläinen, J. (2002). Estimating drainage area-based snow-cover percentages from NOAA AVHRR images. *International Journal of Remote Sensing*, **23**(15):2971–2988.

Metsämäki, S., Vepsäläinen, J., Pulliainen, J., and Sucksdorff, Y. (2002). Improved linear interpolation method for the estimation of snow-covered area from optical data. *Remote Sensing of Environment*, **82**(1):64–78.

Metsämäki, S. J., Anttila, S. T., Markus, H. J., and Vepsäläinen, J. M. (2005). A feasible method for fractional snow cover mapping in boreal zone in boreal zone based on a reflectance model. *Remote Sensing of Environment*, **95**:77–95.

Middleton, W. E. K. and Mungall, A. G. (1952). The luminous directional reflectance of snow. *Appl. Opt.*, **42**(8):572–579.

Mishchenko, M. I. (2008). Multiple scattering, radiative transfer, and weak localization in discrete random media: unified microphysical approach. *Reviews of Geophysics*, **46**(RG2003).

Mishchenko, M. I., Dlugach, J. M., Yanovitskij, E. G., and Zakharova, N. T. (1999). Bidirectional reflectance of flat, optically thick particulate layers: An efficient radiative transfer solution and applications to snow and soil surfaces. *J. Quant. Spectrosc. Radiat. Transfer*, **63**(2-6):409–432.

Mishchenko, M., Hovenier, J., and Travis, L., editors (2000). *Light Scattering by Non-Spherical Particles: Theory, Measurements and Applications*. Academic Press, San Diego, CA.

Moody, E. G., King, M. D., Schaaf, C. B., Hall, D. K., and Platnick, S. (2007). Northern hemisphere five-year average (2000–2004) spectral albedos of surfaces in the presence of snow: Statistics computed from Terra MODIS land products. *Remote Sensing of Environment*, **111**:337–345.

Muinonen, K. O. (2004). Coherent backscattering of light by complex random media of spherical scatterers: numerical solution. *Waves in Random Media*, **14**:365–388.

Muinonen, K., Nousiainen, T., Fast, P., Lumme, K., and Peltoniemi, J. I. (1996). Light scattering by Gaussian random particles: ray optics approximation. *J. Quant. Spectrosc. Radiat. Transfer*, **55**(5):577–602.

Nakaya, U. (1954). *Snow Crystals: Natural and Artificial*. Harvard University Press.

Nicodemus, F. E., Richmond, J. C., Hsia, J. J., Ginsberg, I. W., and Limperis, T. (1977). Geometrical considerations and nomenclature for reflectance. Technical report, Institute for Basic Standards, National Bureau of Standards, Washington DC, USA.

Nolin, A. W. and Dozier, J. (1991). Spectral reflectance of thin snow. In *Proc. of the Fifth International Colloquium on Physical Measurements and Signatures in Remote Sensing*, ESA SP-319, pages 439–442, Courchevel, France.

Nolin, A. W., Dozier, J., and Davis, R. E. (1990). Bidirectional reflectance of optically-thin snow. In *Proc. of the International Geoscience and Remote Sensing Symposium '90*, IEEE 90CH2825-8, pages 1159–1161, College Park, Maryland.

Nolin, A. W. and Dozier, J. (2000). A hyperspectral method for remotely sensing the grain size of snow. *Remote Sensing of Environment*, **74**:207–216.

Nolin, A. W. and Stroeve, J. C. (2001). Validation studies and sensitivity analyses for retrievals of snow albedo from EOS AM-1 instruments. progress report for year 2000-2001. Technical report, National Snow and Ice Data Center, Cooperative Institute for Research in Environmental Sciences.

Nolin, A. W. and Frei, A. (2001). Remote sensing of snow and snow albedo characterization for climate simulations. In Beniston, M. and Verstraete, M. M., editors, *Remote Sensing and Climate Simulations: Synergies and Limitations*, Advances in Global Change Research, pages 159–180. Kluwer Academic Publishers, Dordrecht and Boston.

Nolin, A. W., Steffen, K., and Dozier, J. (1994). Measuring and modeling the bidirectional reflectance of snow. In *Proc. of the International Geoscience and Remote Sensing Symposium '94*, IEEE 94CH3378-7, pages 1919–1921, Pasadena, CA.

Odermatt, D., Schläpfer, D., Lehning, M., Schwikowski, M., Kneubühler, M., and Itten, K. I. (2005). Seasonal study of directional reflectance properties of snow. In *EARSeL eProceedings*, volume 4, pages 203–214.

Painter, T. H. and Dozier, J. (2002). Measurements of the bidirectional reflectance of snow at fine spectral and angular resolution. In *70th Annual Meeting of the Western Snow Conference*, page 70.

Painter, T. H. and Dozier, J. (2004). Measurements of the hemispherical–directional reflectance of snow at fine spectral and angular resolution. *J. Geophys. Res.*, **109**(D18115).

Painter, T. H., Dozier, J., Roberts, D. A., Davis, R. E., and Green, R. O. (2003a). Retrieval of subpixel snow-covered area and grain size from imaging spectrometer data. *Remote Sensing of Environment*, **85**:64–77.

Painter, T. H., Paden, B., and Dozier, J. (2003b). Automated spectro-goniometer: A spherical robot for the field measurement of the directional reflectance of snow. *Review of Scientific Instruments*, **74**(12):5179–5188.

Painter, T. H., Rittger, K., McKenzie, C., Slaughter, P., Davis, R. E., and Dozier, J. (2009). Retrieval of subpixel snow covered area, grain size, and albedo from MODIS. *Remote Sensing of Environment*, **113**:868–879.

Peltoniemi, J. I. (2007). Spectropolarized ray-tracing simulations in densely packed particulate medium. *J. Quant. Spectrosc. Radiat. Transfer*, **108**(2):180–196.

Peltoniemi, J. I. and Lumme, K. (1992). Light scattering by closely packed particulate media. *J. Opt. Soc. Am. A*, **9**(8):1320–1326.

Peltoniemi, J. I., Lumme, K., Muinonen, K., and Irvine, W. M. (1989). Scattering of light by stochastically rough particles. *Appl. Opt.*, **28**(19):4088–4095.

Peltoniemi, J., Viljakainen, M., Levander, T., Hiltunen, S., Piironen, J., and Näränen, J. (2002). Another field goniometer. In *Abstracts of Third International Workshop on Multiangular Measurements and Models*, page 42, Steamboat Springs, Colorado. Poster.

Peltoniemi, J., Kaasalainen, S., Näränen, J., Matikainen, L., and Piironen, J. (2005a). Measurement of directional and spectral signatures of light reflectance by snow. *IEEE Transactions on Geoscience and Remote Sensing*, **43**(10):2294–2304.

Peltoniemi, J., Kaasalainen, S., Näränen, J., Rautiainen, M., Stenberg, P., Smolander, H., Smolander, S., and Voipio, P. (2005b). BRDF measurement of understory vegetation in pine forests: dwarf shrubs, lichen and moss. *Remote Sensing of Environment*, **94**:343–354.

Peltoniemi, J., Piironen, J., Näränen, J., Suomalainen, J., Kuittinen, R., Honkavaara, E., and Markelin, L. (2007). Bidirectional reflectance spectrometry of gravel at the

Sjökulla test field. *ISPRS Journal of Photogrammetry and Remote Sensing*, **62**(6):434–446.

Peltoniemi, J., Hakala, T., Suomalainen, J., and Puttonen, E. (2009). Polarised bidirectional reflectance factor measurements from snow, soil and gravel. *J. Quant. Spectrosc. Radiat. Transfer*, **110**:1940–1953.

Perovich, D. K. (1990). Theoretical estimates of light reflection and transmission by spatially complex and temporally varying sea ice covers. *J. Geophys. Res.*, **95**(C6):9557–9567.

Perovich, D. K. (1994). Light reflection from sea ice during the onset of melt. *J. Geophys. Res.*, **99**(C2):3351–3360.

Pirazzini, R. (2008). *Factors controlling the surface energy budget over snow and ice*. PhD thesis, University of Helsinki, Faculty of Science, Department of Physics, Division of Atmospheric Sciences and Geophysics.

Puttonen, E., Suomalainen, J., Hakala, T., and Peltoniemi, J. (2009). Measurement of reflectance properties of asphalt surfaces and their usability as reference targets for aerial photos. *IEEE Transactions on Geoscience and Remote Sensing*, **47**(7):2330–2339.

Roesch, A., Wild, M., and Ohmura, A. (2001). Snow cover fraction in a general circulation model. In Beniston, M. and Verstraete, M. M., editors, *Remote Sensing and Climate Simulations: Synergies and Limitations*, Advances in Global Change Research, pages 203–232. Kluwer Academic Publishers, Dordrecht and Boston.

Salminen, M., Pulliainen, J., Metsämäki, S., Kontu, A., and Suokanerva, H. (2009). The behaviour of snow and snow-free surface reflectance in boreal forests: Implications to the performance of snow covered area monitoring. *Remote Sensing of Environment*, **113**:907–918.

Schopfer, J. T. (2008). *Spectrodirectional ground–based remote sensing using dual–view goniometry*. PhD thesis, University of Zurich.

Solberg, R., Hiltbrunner, D., Koskinen, J., Guneriussen, T., Rautiainen, K., and Hallikainen, M. (1997). Snow algorithms and products, review and recommendations for research and development. Report from SNOWTOOLS WP 410 924, Norwegian Computing Center, Oslo.

Steffen, K. (1987). Bidirectional reflectance of snow at 500–600 nm. In *Large Scale Effects of Seasonal Snow Cover (Proceedings of the Vancouver Symposium)*, number 166 in IAHS Publ.

Stroeve, J., Box, J. E., Gao, F., Liang, S., Nolin, A., and Schaaf, C. (2005). Accuracy assessment of the MODIS 16-day albedo product for snow: comparisons with Greenland in situ measurements. *Remote Sensing of Environment*, **94**:46–60.

Stroeve, J. C., Box, J. E., and Haran, T. (2006). Evaluation of the MODIS (MOD10A1) daily snow albedo product over the Greenland ice sheet. *Remote Sensing of Environment*, **105**:155–171.

Suomalainen, J., Hakala, T., Peltoniemi, J., and Puttonen, E. (2009a). Polarised multiangular reflectance measurements using Finnish Geodetic Institute field goniospectrometer. *Sensors*, **9**(5):3891–3907.

Suomalainen, J., Hakala, T., Puttonen, E., and Peltoniemi, J. (2009b). Polarised bidirectional reflectance factor measurements from vegetated land surfaces. *J. Quant. Spectrosc. Radiat. Transfer*, **110**:1044–1056.

Tanikawa, T., Aoki, T., Hori, M., Hachikubo, A., and Aniya, M. (2006). Snow bidirectional reflectance model using non-spherical snow particles and its validation with field measurements. In *EARSeL eProceedings*, volume 5, pages 137–145.

Tedesco, M. and Kokhanovsky, A. (2007). The semi-analytic snow retrieval algorithm and its application to MODIS data. *Remote Sensing of Environment*, **111**:228–241.

Warren, S. G. (1982). Optical properties of snow. *Reviews of Geophysics and Space Physics*, **20**(1):67–89.

Warren, S., B.-C.Gao, and Wiscombe, W. (1995). ftp//climate.gsfc.nasa.gov/pub/wiscombe/Refrac_Index.

Warren, S., Brandt, R., and Hinton, P. O. (1998). Effect of surface roughness on bidirectional reflectance of Antarctic snow. *J. Geophys. Res. (Planets)*, **103**(E11):25789–25807.

Widen, N. (2000). A description of the new FGI goniometer and quality analysis of experimental data. *Photogrammetric Journal of Finland*, **17**(1):28.

Wiscombe, W. and Warren, S. (1980). A model for spectral albedo of snow I: pure snow. *J. Atmos. Sci.*, **37**(12):2712–2733.

10 Simulation and modeling of light scattering in paper and print applications

Per Edström

10.1 Introduction

When developing and applying models to light scattering problems, things usually turn very mathematical. This is all in good order, but it may also be a hindrance for a broader audience to gain insight into the overall issues. This chapter aims at discussing a range of light scattering simulation and modeling issues with a minimum of mathematics involved, and with the specific perspective of paper and printing industry applications. Shorter sections of mathematical content are included, but the mathematically interested reader is here pointed to selected references and other chapters in this volume.

Section 10.2 gives an overview of the standardized use of models and measurements in these industries, and includes some drawbacks and needs for newer models in these applications. Section 10.3 discusses the benefits of radiative transfer and Monte Carlo models in paper and printing industry applications, and covers the impact on measurement systems and industry standards. Finally, section 10.4 sums up current challenges and future work in this application area of light scattering.

10.2 Current industrial use of light scattering models

The paper and print industries have been using two-flux light scattering models since the 1930s. The main purpose has been to describe and predict quantities like reflectance, transmittance, opacity and brightness, but also derived subjective quantities like whiteness and various aspects of color. This is done for papers containing different pulps and fillers, or papers with multilayer structures, e.g. coated papers. Simulation and modeling is important for prediction in the development of new products, and it is used for quality control in production (mainly printing industry), and for data exchange in trading situations (mainly paper industry).

10.2.1 Standardized use of Kubelka–Munk

As is well known, the Kubelka–Munk light scattering and light absorption coefficients (s and k) are widely used in the pulp and paper industry in applications

ranging from research projects to practical problems in paper mills. When assessing optical properties of pulp and paper, it is primarily the diffuse reflection that is used. A series of ISO standards [1–4] state how measurements should be made, and how the Kubelka–Munk [5–7] theory should be used for interpretation of measurement data. Regarding reflectance measurements, what is usually meant is the diffuse reflectance factor, which is defined as the ratio of the radiation reflected by a body to that reflected by the perfect reflecting diffusor under the same conditions of illumination and detection. The perfect reflecting diffusor is not a real medium, but a basic concept in paper optics.

10.2.1.1 Short formulation of the Kubelka–Munk theory

There are other theories that in greater detail describe the interaction of light with matter, but the Kubelka–Munk theory is simple enough for widespread use, yet accurate enough for many industrial applications. The Kubelka–Munk equations are approximately valid when illumination and measurement take place in diffuse conditions, and when gloss effects (i.e. directed reflectance from a surface) are eliminated.

Kubelka–Munk is a two-flux model, and diffuse light is considered to be travelling in the lower or upper hemisphere with the respective intensities i and j. A plane-parallel medium of thickness w is considered, and x is the distance from the background (or the bottom of the medium). Thus, a thin layer dx is illuminated with intensity $i(x)$ from above and $j(x)$ from below. The model uses the two phenomenological parameters s and k for scattering and absorption, and when light passes through the layer dx, i and j are reduced by $(s+k)i\,dx$ and $(s+k)j\,dx$ respectively, due to scattering and absorption. There is also the addition of $sj\,dx$ and $si\,dx$ respectively, due to scattering from the other direction. The fundamental equations of Kubelka–Munk are thus

$$\begin{cases} -di = -(s+k)i\,dx + sj\,dx \\ \quad dj = -(s+k)j\,dx + si\,dx. \end{cases} \tag{10.1}$$

The different sign in the left-hand side of these equations is due to the opposite directions of propagation of the respective light fluxes.

It is customary in the paper industry to let x have the unit kg/m^2, and thereby s and k have the unit m^2/kg. In line with this, w is then called grammage. This was suggested by van den Akker [8], who showed that the original Kubelka–Munk differential equations remain unchanged under this transformation. He proposed this based on the fact that in practical application of the Kubelka–Munk model, the thickness of a paper may change significantly without affecting the reflectance, suggesting that the light scattering remains unchanged, which naturally is also true for the grammage. The use of grammage is now common practice in paper-related applications.

The Kubelka–Munk equations (10.1) are simple enough to have a solution that can be analytically expressed in both a forward and an inverse direction, which of course is part of the explanation of its widespread industrial use. If a paper sample is illuminated with light of intensity I and reflected light is measured to have intensity J, the reflectance is given by

$$R = J/I. \tag{10.2}$$

The reflectance of the background of the sample, denoted by R_g, of course has an influence on the result, and it is customary to make measurements for R_0, when $R_g = 0$, and for R_∞, for an opaque pad of samples so that R_g has no influence. Some useful relations can then be stated as

$$s = \frac{1}{w\left(\frac{1}{R_\infty} - R_\infty\right)} \ln\left(\frac{(1 - R_0 R_\infty)R_\infty}{R_\infty - R_0}\right), \tag{10.3}$$

$$k = \frac{s(1 - R_\infty)^2}{2R_\infty}, \tag{10.4}$$

$$R_\infty = 1 + \frac{k}{s} - \sqrt{\left(\frac{k}{s}\right)^2 + 2\frac{k}{s}} \tag{10.5}$$

and

$$R_0 = \frac{R_\infty \exp\left[sw\left(\frac{1}{R_\infty} - R_\infty\right)\right] - R_\infty}{\exp\left[sw\left(\frac{1}{R_\infty} - R_\infty\right)\right] - R_\infty^2}. \tag{10.6}$$

Opacity is then defined as

$$Op = 100\frac{R_0}{R_\infty}. \tag{10.7}$$

10.2.1.2 Optical geometry of measurement devices

The pulp and paper industry primarily uses an instrument with d/0 geometry, which means diffuse illumination and detection perpendicular to the sample. Diffuse illumination is usually accomplished with an integrating sphere, coated on the inside with white barium sulphate. The sample and detector are screened from direct illumination from the light source, and directed reflection is eliminated with a gloss trap consisting of a black screen around the detector.

For color measurements in the printing industry, 45/0 instruments are widely used. There, the sample is illuminated at an angle of 45 degrees, and detection is perpendicular to the sample. In this way, gloss is automatically eliminated, except for rough surfaces.

Optical measurement devices differ in more ways that just optical geometry. Paper industry devices have a traceable calibration procedure and measure on a relatively large area, while printing industry devices just use a relative calibration and measures a smaller area. There are also differences in the size of the measured area in relation to the illuminated one, and in the spectral content of the illumination. Differences in UV content of the illumination has a large impact for samples that fluoresce. Instruments can also be equipped with various filters. There are thus several reasons that identical samples get different results when measured in different instruments, and these differences can well be large enough to cause waste in production.

10.2.2 Deficiencies of Kubelka–Munk

Several limitations for the Kubelka–Munk model have been reported, e.g. concerning dependencies between the scattering and absorption parameters s and k for translucent or strongly absorbing media [8–13], and attempts have been made to attribute some of this behavior to intrinsic errors of the Kubelka–Munk model [14–18], or to phenomena not included in it. Despite these limitations, the Kubelka–Munk model is in widespread use for light scattering calculations in the pulp, paper and printing industries, probably due to its explicit form and ease of use. These are also sufficient reasons for continued industrial usage where the accuracy is sufficient, and where there are no reported limitations. However, new solution methods with better accuracy and a larger range of applicability should be considered in many cases.

The interpretation of Kubelka–Munk coefficients in terms of structural features is not always obvious. One example is their behavior for highly absorbing samples, e.g. heavily dyed papers. It is observed that the apparent scattering coefficient is reduced after dyeing at wavelengths corresponding to the absorption of the dyeing colorant [12]. Since scattering is a feature of the structure of the sample, and dyeing is not supposed to change the structure, this decrease is regarded as physically incorrect. It has been shown [13, 17] that this decrease in part can be explained by the inability of the Kubelka–Munk model to account for the anisotropy of scattered light that is introduced by the higher absorption.

The anisotropy of the scattered light in real situations can often be large, and can significantly influence the measurements. This holds even for situations with isotropic single-scattering process or diffuse illumination, which might be non-intuitive. The standardized use of the Kubelka–Munk model for interpretation of measurement data implies a far too idealized view of the measurement situation. This is especially noticeable for optically thin and highly absorbing samples [13]. Since the Kubelka–Munk model cannot accommodate for anisotropy, an angle-resolved model is needed for measurement data interpretation in such cases, which are not infrequent in practice.

Although numerical values of the asymmetry factor for paper samples are not frequent in the literature, the few notes that are made suggest that the single-scattering process in paper media is not at all isotropic [19]. This means that no matter how one tries to create diffuse illumination, the Kubelka–Munk model will not handle the basic scattering process correctly. This also calls for an angle-resolved model, which would accommodate for any single-scattering process as well as for the following anisotropy of light intensities.

There are also a few areas where it has always been clear that Kubelka–Munk does not suffice, and where new and higher demands now drive the need for new models. This is briefly covered and exemplified in subsection 10.2.4.

10.2.2.1 Some illustrated examples

The DORT2002 model [20] can be used to study how the d/0 instrument reflectance factor, $R_{d/0}$, corresponds to the total reflectance for a chosen parameter space. The comparison between $R_{d/0}$ and the total reflectance has been done for a parameter

space $0 < \tau < 20$ and $0 < \omega_0 < 1$, τ being the optical thickness and ω_0 the single scattering albedo, and the result is shown in Fig. 10.1. The simulation shows that $R_{d/0}$ and the total reflectance as calculated with DORT2002, $R_{DORT2002}$, diverge for optically thin and strongly absorbing media, with the dependence on optical thickness being very significant. If the reflected light were isotropic, the d/0 reflectance factor and the total reflectance would coincide, but they are seen to differ and the difference depends strongly on medium parameters. This is a direct consequence of the anisotropic nature of the intensity of the reflected light. This effect can also be illustrated by results of asymptotic radiative transfer. Kokhanovsky [21–23] gives some approximations for very thick and highly scattering media. With

$$y = 4\sqrt{\frac{1-\omega_0}{3(1-g\omega_0)}}, \tag{10.8}$$

the spherical albedo is given as

$$r_{s\infty}(\mu) = \exp(-y), \tag{10.9}$$

and the plane albedo as

$$r_{p\infty}(\mu) = \exp(-3y(1+2\mu)/7). \tag{10.10}$$

The spherical albedo is the hemispherical reflectance for diffuse illumination, and corresponds to total reflectance above. The plane albedo is the hemispherical reflectance for collimated illumination, but due to the reciprocity principle it also holds for diffuse illumination and directed observation (μ is cosine of this direction). This means that $R_{d/0}$ above can be approximated by $\exp(-9y/7)$. It is also explicitly clear from Eq. (10.10) that the reflected light shows directional dependence, with less light reflected in normal directions.

This explains the shape of Fig. 10.1. The deviation of $R_{d/0}$ from the total reflectance is, as can be seen from Eq. (10.10), due to a higher reflectance in larger polar angles. The size of the deviation also depends on medium thickness and absorption. The higher the absorption and the thinner the medium, the more light will be scattered into larger polar angles, and the d/0 instrument cannot detect this light. For a non-absorbing medium with $g = 0$, the reflectance approaches the perfectly diffuse when the medium gets infinitely thick.

Simulations also show that the reflectance from a medium changes in a characteristic way when the illumination is altered. For example, the $BRDF$ is changed when a sample is illuminated diffusely and with normally incident light respectively. When the medium parameters are put to those intuitively believed to correspond to the perfect diffusor, i.e. $\omega_0 = 1$, $g = 0$ and $\tau \to \infty$, the supposed perfect diffusor reflects diffusely only if the illumination is diffuse but the reflectance is anisotropic for normally incident light. In practice this will be the case for every supposed perfect diffusor involving bulk scattering, since it is adequately described by the radiative transfer equation. But, if the diffusor were to be constructed from some material only involving surface scattering, i.e. with no light penetrating the medium, the phenomenon could be avoided. It is thus a fundamental feature of every material adequately described by radiative transfer theory that it can never reflect light as a Lambertian surface independently of the illumination conditions.

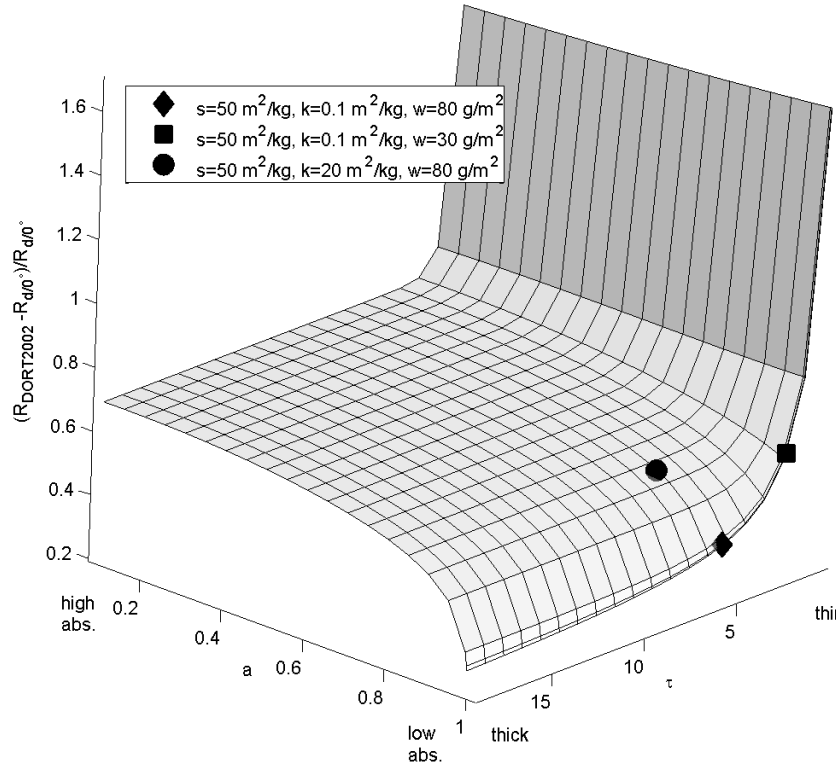


Fig. 10.1. The relative difference between the d/0 reflectance factor and the total reflectance as simulated with DORT2002. It can be seen that, if $R_{d/0}$ is interpreted as total reflectance, the true total reflectance, here denoted $R_{DORT2002}$, is underestimated and that the difference depends strongly on medium parameters. The difference increases with absorption and increases very rapidly when the optical thickness decreases. For intensely scattering and optically thick media it approaches zero. The asymmetry factor g has been put to zero to generate this plot. The coordinates for three Kubelka–Munk parameter setups have been included for reference.

Figs 10.2 and 10.3 show how the intensity of the reflected light is actually distributed in an ideal measurement device of 45/0 and d/0 type, respectively. Both instrument types measure in the normal direction only, and the following Kubelka–Munk calculations therefore assume a perfectly diffuse distribution (inner gray half sphere). It is obvious that this differs greatly from the actual distribution (outer white shape), and this yields errors even for the situation with perfectly isotropic single scattering that the system is designed for (topmost panes). For a more anisotropic single scattering process the error increases, and in all cases the total amount of reflected light is underestimated. In the 45/0 case, the highly asymmetric distribution should be noted, which obviously differs greatly from where Kubelka–Munk should be used. Although numerical values of asymmetry factor for paper are scarcely reported, it should be noted that values in the range 0.3–0.8 are common, which corresponds to the lower panes of Figs 10.2 and 10.3.

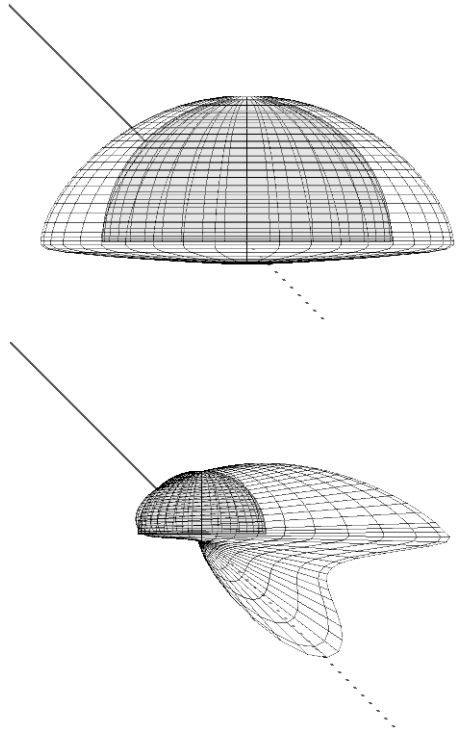


Fig. 10.2. Anisotropic $BRDF$ from sample ($\omega_0 = 0.9$, $\tau = 5.5$) measured in 45/0 instrument geometry for different asymmetry factors ($g = 0$ at top and $g = 0.65$ at bottom), as simulated with DORT2002 and presented in a relative scale. Note that the instrument measures in the normal direction, and therefore incorrectly must assume the reflectance to be the same in all other directions (the gray half sphere). The instrument clearly underestimates the reflectance. Incident beam indicated as solid line.

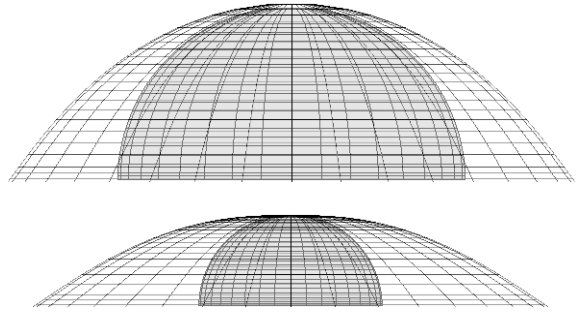


Fig. 10.3. Anisotropic $BRDF$ from sample ($\omega_0 = 0.9$, $\tau = 5.5$) measured in d/0 instrument geometry for different asymmetry factors ($g = 0$ at top and $g = 0.65$ at bottom), as simulated with DORT2002 and presented in a relative scale. Note that the instrument measures in the normal direction, and therefore incorrectly must assume the reflectance to be the same in all other directions (the gray half sphere). The instrument clearly underestimates the reflectance.

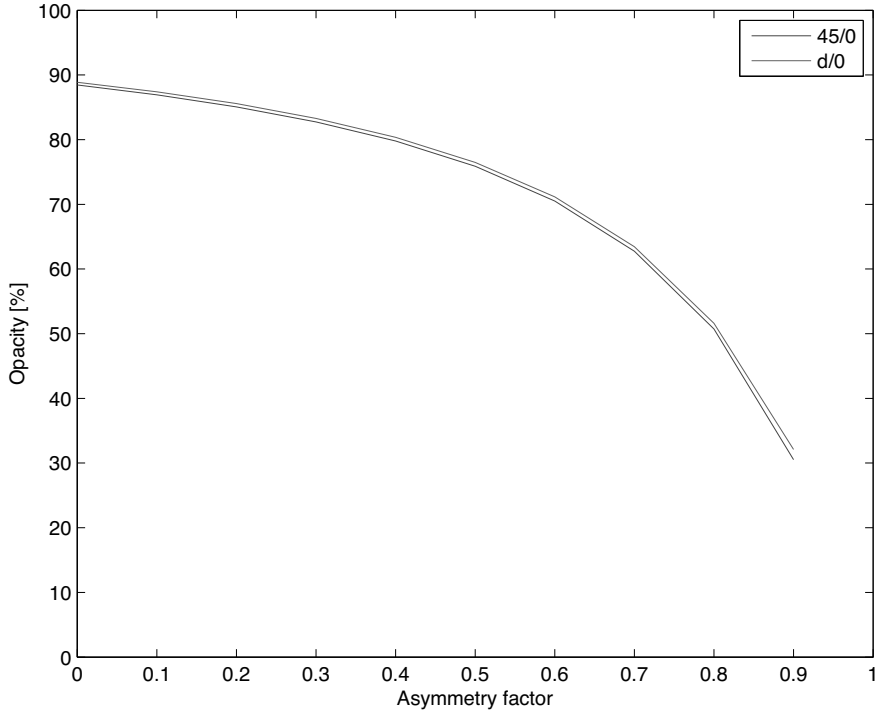


Fig. 10.4. Opacity of sample ($\omega_0 = 0.99$, $\tau = 6$) as calculated with DORT2002 in 45/0 and d/0 instrument geometries for different asymmetry factors. Note the decrease in opacity for large asymmetry factors.

The opacity (defined in Eq. (10.7)) and the reflectance factor over black background were calculated with DORT2002 for identical samples of finite thickness, but with varying asymmetry factor. The simulations mimic the standardized d/0 and 45/0 measurement situations, and are illustrated in Figs 10.4 and 10.5. It is clear how these quantities are markedly affected by the asymmetry factor, and the behavior approximately follows the $\exp(-A(1-g)^{-1/2})$ law, thus decreasing for larger asymmetry factor.

It has, as mentioned above, been reported on apparent dependencies between the Kubelka–Munk scattering and absorption coefficients for dyed paper samples. Light scattering is largely due to the structure of the sample, which is not affected by the addition of dye. Therefore, identical paper samples with increasing amount of dye should have nearly similar scattering coefficients. However, it is reported how the scattering coefficient decreases in the region of absorption, and this has been interpreted as an intrinsic error of the Kubelka–Munk model. Fig. 10.6 illustrates measurements of this for dyed paper samples with fillers.

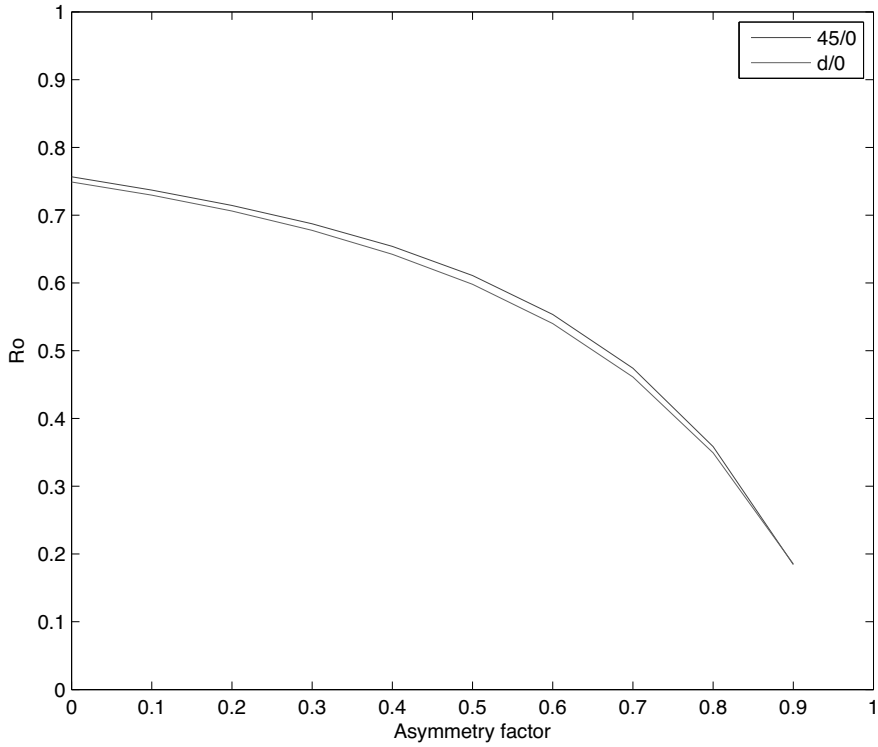


Fig. 10.5. Reflectance factor over black background of sample ($\omega_0 = 0.99$, $\tau = 6$) as calculated with DORT2002 in 45/0 and d/0 instrument geometries for different asymmetry factors. The anisotropy approximately follows the $\exp(-A(1-g)^{-1/2})$ law, although the sample thickness is finite. Note the decrease in reflectance factor for large asymmetry factors.

10.2.3 Suggested extensions to Kubelka–Munk

The Kubelka–Munk model has been modified and extended for different purposes in a variety of ways [24]. Most suggestions are, however, of limited generality, although they yield somewhat improved results for certain specific purposes. This is fine, as long as the new models are used only for that given purpose. One example is given by Murphy [25], where one finds an interesting method for estimating optical properties of coatings, based on an extended Kubelka–Munk approach. Kokhanovsky [26] introduces a modified Kubelka–Munk approximation for optically thick media and deduces analytical relations to asymptotic radiative transfer theory for this case. Yet another alternative is to suggest analytical approximations to radiative transfer theory for specific cases, without even relating to Kubelka–Munk. An example of this is given by Kokhanovsky and Hopkinson [27], who deduce an analytical approximation to radiative transfer theory for optically thick and weakly absorbing media, and thereby relate material parameters to diffuse reflectance measurements for this case.

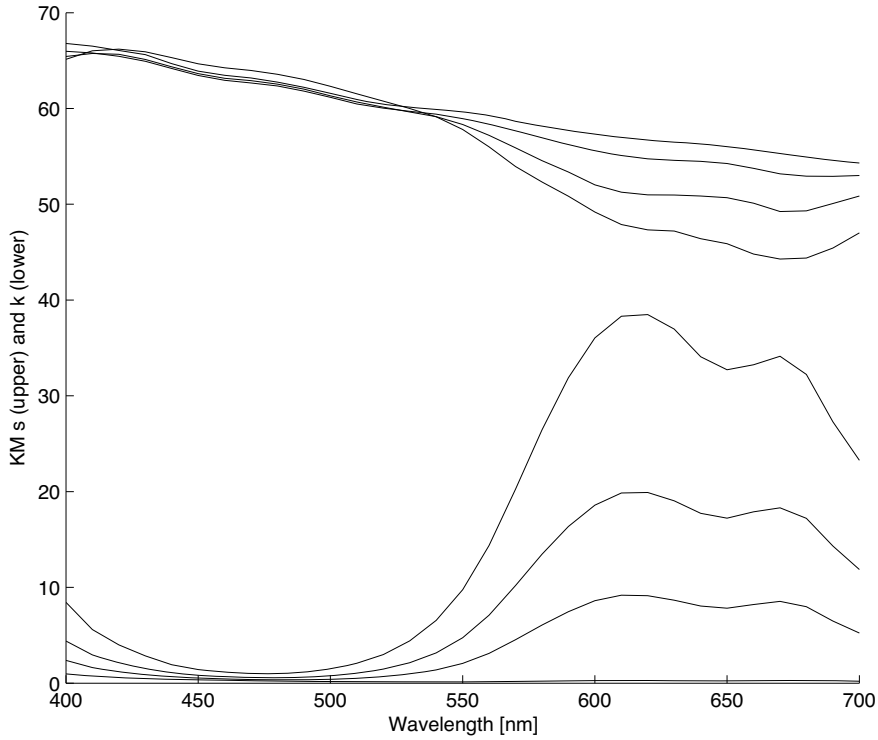


Fig. 10.6. Anomalous parameter dependencies for dyed paper samples with fillers, when measured according to paper industry standards and calculated using Kubelka–Munk. Upper curves are KM scattering coefficient s, and lower curves are KM absorption coefficient k. The decrease in s in the region of absorption has been interpreted as an intrinsic error of KM theory. Experimental details can be found in [38].

The Kubelka–Munk coefficients s and k have no direct physical meaning on their own, and should only be interpreted within the Kubelka–Munk model; they do not represent anything physically objective outside the Kubelka–Munk model. This is contrary to the usual formulation of the radiative transfer problem, where the scattering and absorption parameters are related to the mean free path of the medium, and are thus model- and geometry-independent. They can therefore be given a physically objective interpretation, which is a desirable feature for modeling. There are approximate relations between the Kubelka–Munk coefficients and the physically objective absorption and scattering coefficients (σ_a and σ_s), like

$$s = \frac{3}{4}\sigma_s(1-g), \tag{10.11}$$

$$k = 2\sigma_a, \tag{10.12}$$

suggested by Mudgett and Richards [28, 29] and complemented by van de Hulst [30]. No exact translation can exist, however, since Kubelka–Munk is incommensurable with higher-order models; Kubelka–Munk is fundamentally simpler and a translation to higher-order models could never be complete. Indeed, the existence

of a complete translation would actually imply that the higher-order model was equivalent to the simpler Kubelka–Munk, which would be a contradiction in terms. Instead, relations like these should be regarded as the first term of some series expansion. There is, however, a reduced equivalence for infinitely thick non-absorbing samples, and for that case it is possible to deduce a complete translation.

Some authors have tried to improve the Kubelka–Munk model as such, not for a specific purpose only, but more as an attempt to actually do a better job than Kubelka and Munk in that same setting. Yang and co-workers recently proposed a revised Kubelka–Munk theory [31–34], arguing that there was an oversight in the derivation of the original Kubelka–Munk theory. This was debated, however, by Edström [35], who also discussed different modeling strategies and their feasibility. Edström argues that Kubelka–Munk is as 'exact' as it can be within the two-flux approximation. There is therefore no such general two-flux improvement to be made, and those kinds of extensions of Kubelka–Munk are of little value. Modifications for specific purposes are still relevant, however, as mentioned above.

Some authors use random walk discussions and various reasoning from statistical mechanics (e.g. [36]), which is fine, but others are dangerously close to mixing finite and infinitesimal properties in an unsatisfactory way.

10.2.4 New and higher demands drive the need for new models

Measurement methods have been developed independently in the paper and printing industries, and are today incompatible. To face market challenges, better models are required for specifying and communicating optical properties, or inefficiencies will hinder competitiveness. Accurate color reproduction is difficult since papers, inks, and printing methods usually differ between the proof and the actual production. The proof is usually printed on another paper and with a different printing process. Different illumination conditions between proof and print may give a very different perception of the colors. This proof-to-production problem causes waste, and this step is made even harder if fluorescing agents are used in the paper.

Apart from the obvious need for harmonization of standards and calibration routines, better models are called for to accommodate accurate treatment of fluorescence and a transfer to physically objective material parameters. The latter is needed for accurate data exchange between different measurement situations, and requires better models for interpretation of measurement data.

There has for a long time been a broad understanding that gloss is not handled by Kubelka–Munk, and this has led to various attempts to eliminate gloss in measurement situations. This is not always possible, however, and gloss is still present in real situations. There is therefore a need for models that can handle gloss effects, and thereby, for example, distinguish between surface and bulk contributions to reflectance measurements.

It is well known that the Kubelka–Munk model has limitations, but it is still also used outside its range of applicability, probably because of its ease of use and the lack of accessible alternatives. For very thin samples and for highly absorbing samples more accurate models are needed. Apart from solving the full radiative transfer equation, there are different approximations. Kokhanovsky [23] states an

expression for the reflection function for high absorption as

$$R_\infty(\xi,\eta,\varphi) = \frac{\omega_0 p(\cos\Theta)}{4(\xi+\eta)(1-\omega_0 F)}, \tag{10.13}$$

where ξ and η are cosine of polar angle for incident and observation directions, and $F \approx 1$.

Within the paper and printing industries, there is but an emerging understanding that anisotropy in the light distribution affects measurements, and this far more than the common idealized view of the measurement situation. The need for angle-resolved models will grow together with this understanding.

There is an increasing desire for deeper knowledge of the effects of gloss, fluorescence, anisotropy and more, not just to be able to measure and interpret data, but also to make accurate predictions and thereby include such phenomena in the optical design of new products. This ability will give a competitive advantage.

10.3 Benefits of newer models

There are numerous models in the light scattering community that could solve selections of paper and printing industry problems, but most are not sufficiently accessible or specific for industrial use. There are, however, both radiative transfer and Monte Carlo models that are adapted to this area of application, and different benefits can be identified.

10.3.1 Radiative transfer modeling

The Kubelka–Munk model is actually itself a radiative transfer model, although the simplest form – a two-flux model. The natural next step is of course to use a many-flux model or, which might be a better term, an angle-resolved model. One such adapted model in recent use is Dort2002 [20], which is a discrete ordinate method. Dort2002 has also been extensively studied regarding numerical performance in both a forward [37] and an inverse problem setting [38, 39], and it has been successfully applied to real paper industry problems [13, 17, 40].

One obvious advantage in model development of going from a two-flux to a many-flux is that this is a true generalization. Therefore, existing measurement devices and results can still be used, and nothing needs to be discarded. Since the same setting can be handled, it simply serves the possibility of more accurate interpretation of measurement data. There are, however, a number of more profound advantages.

Instead of using the phenomenological parameters s and k, it is possible to use the physically objective parameters σ_s and σ_a, and the phase function p (see next section for definition). Since s and k get their meaning through the Kubelka–Munk model, they have no objective meaning outside, and they are both model- and geometry-dependent. Physically objective material parameters are independent of model and measurement geometry, and are thus better suited for data exchange.

An angle-resolved model also accounts for anisotropy through the phase function p, which is often characterized through the asymmetry factor g. This makes it

possible to account for differences in measurement geometry, which is ignored in the over-idealized view of the measurement situation that is necessary when using Kubelka–Munk. Since anisotropy is ever-present, the ability to include it in a model also makes it possible to understand the phenomenon, but more importantly to use it as a design parameter in new products. To be able to take full advantage of this, however, angle-resolved measurements need to be used.

10.3.1.1 Radiative transfer forward problem formulation

For an ideally reflecting medium, all incoming light is specularly reflected at the surface. For a turbid medium, transmission as well as absorption and multiple scattering inside the medium have to be taken into consideration. In paper applications, the problem is studied in a plane-parallel geometry, where the horizontal extension of the medium is assumed to be large enough to give no boundary effects at the sides. The boundary conditions at the top and bottom, including illumination, are assumed to be time- and space-independent. The radiation is assumed to be monochromatic, or confined to a narrow wavelength range. The scattering is assumed to be without change in frequency between incoming and outgoing radiation. The medium is treated as a continuum of scattering and absorption sites. Polarization effects are often ignored, hence using only the first component of the Stokes 4-vector. What is left is then a scalar intensity, which is the variable to solve for.

The energy flow is considered as non-interacting beams of radiation in all directions. This makes it possible to treat the beams separately. The intensity, I, of the radiation is always considered to be positive. When radiation traverses a finite thickness ds of the medium in its direction of propagation, a fraction is extinct due to absorption and scattering. The intensity then becomes $I + dI$, and the extinction coefficient is defined as

$$\sigma_e = -\frac{dI}{I\,ds}. \tag{10.14}$$

The extinction coefficient can be separated into two parts, the absorption and scattering coefficients σ_a and σ_s, corresponding to the different origins of the extinction. The coefficients are related through $\sigma_e = \sigma_a + \sigma_s$, and the single scattering albedo, which is the probability for scattering given an extinction event, and is then defined as

$$\omega_0 = \frac{\sigma_s}{\sigma_e} = \frac{\sigma_s}{\sigma_a + \sigma_s}. \tag{10.15}$$

The phase function, p, specifies the angular distribution of the scattering process. If the phase function is normalized by

$$\int_0^{2\pi} \int_0^{\pi} \sin\theta \frac{p(\theta', \varphi'; \theta, \varphi)}{4\pi}\, d\theta\, d\varphi = 1, \tag{10.16}$$

where θ and φ are the polar and azimuthal angle coordinates of spherical geometry for the direction of the radiation (primed arguments correspond to incident radiation), this can be given a probabilistic interpretation. When radiation in the direction (θ', φ') is scattered, the probability that it is scattered into the cone of solid angle $d\theta d\varphi$ centered on the direction (θ, φ) is

$$\frac{p(\theta', \varphi'; \theta, \varphi)\, d\theta\, d\varphi}{4\pi}. \tag{10.17}$$

Different phase functions have been proposed to physically describe various types of scattering. Among the best known are the phase functions given by Rayleigh [41] and Mie [42]. A very commonly used one-parameter analytical approximation of a real phase function is the Henyey–Greenstein [43] phase function. It is given by

$$p(\cos\Theta) = \frac{1 - g^2}{(1 + g^2 - 2g\,\cos\Theta)^{3/2}}, \tag{10.18}$$

where Θ is the total scattering angle. It is related to the angular coordinates of propagation through the cosine law of spherical geometry as

$$\cos\Theta = \cos\theta'\,\cos\theta + \sin\theta'\,\sin\theta\,\cos(\varphi' - \varphi). \tag{10.19}$$

The coefficients for Legendre polynomial expansion (also referred to as moments) of the Henyey–Greenstein phase function are simply $\chi_l = g^l$. The parameter g is called the asymmetry factor and controls the scattering pattern, ranging from complete forward scattering ($g = 1$) over isotropic scattering ($g = 0$) to complete backward scattering ($g = -1$). An illustration of this is shown in Fig. 10.7.

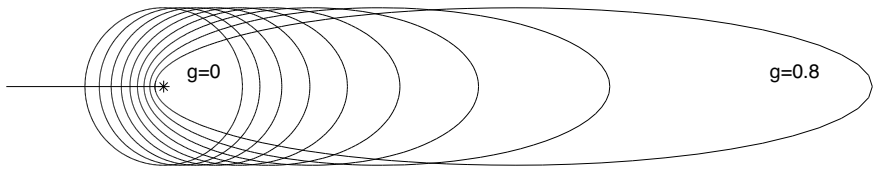


Fig. 10.7. The probability for scattering directions using the Henyey–Greenstein phase function and different values for the asymmetry factor g. Light is incident from the left and the scattering event occurs at the point marked with an asterisk.

In the plane-parallel geometry, distances are measured normal to the surface of the medium. This coincides with the z-axis in a Cartesian coordinate system if the surface is placed in the x-y-plane, and therefore $dz = ds\cos\theta$. The optical depth is measured from the top surface and down, and is thus defined as

$$\tau(z) = \int_z^\infty \sigma_e\, dz'. \tag{10.20}$$

Introducing $u = \cos\theta$, gives $d\tau = -\sigma_e u\, ds$, and Chandrasekhar [44, eq. I.71] states the equation of radiative transfer for a scattering plane-parallel medium as

$$u\frac{dI(\tau, u, \varphi)}{d\tau} = I(\tau, u, \varphi) - \frac{\omega_0}{4\pi}\int_0^{2\pi}\int_{-1}^{1} p(u', \varphi'; u, \varphi) I(\tau, u', \varphi')\, du'\, d\varphi'. \tag{10.21}$$

The integral term is a source function, which gives the intensity scattered from all incoming directions at a point to a specified direction, and it is evident how the phase function acts as a kernel in this integral. It is possible to add a term

for emission, like fluorescence or thermal emission, to the source function if the emission is inside the wavelength range of interest. To perform the coupling of the intensities of the different wavelengths associated with fluorescence, an outer loop over wavelengths will be needed in a solution procedure.

The integro-differential equation of radiative transfer (10.21), can be treated in various ways, but solution methods will not be given here. Suffice it to say that discrete ordinate methods are common (see other chapters in this book), and they generally work as follows. Fourier analysis gives a system of equations, which are discretized using numerical quadrature. The problem can then be transferred to a problem on eigenvalues of matrices. Boundary and continuity conditions are imposed, and the computed intensity is extended from the quadrature points to the entire interval through interpolation formulas.

10.3.1.2 Radiative transfer inverse problem formulation

Using reflectance measurements to solve for only scattering and absorption parameters is rather straightforward, and many methods do this in an optimization procedure with an approximate forward problem like the diffusion approximation, the Eddington approximation or the Kubelka–Munk model. Methods for also estimating the asymmetry factor g are scarce, and even fewer are efficient and accurate. Most efforts come from medical applications, while the industrial side has shown less interest so far. Van Gemert and Star [45] presented an attempt, but it included approximations and boundary conditions that give it limited practical use. Prahl et al. [46] introduced the inverse adding-doubling method, but reported problems in finding suitable initial guess values. One recent interesting approach of high accuracy was presented by Joshi et al. [47], but it suffers from long computation times. Edström [38] recently presented a two-phase method that uses only simple implementations of standard optimization methods for this parameter estimation problem, and still achieved good efficiency. The successful recovery of σ_s, σ_a and g by this two-phase method was illustrated by application to relevant paper industry problems. Of course, for thick samples, only σ_a and $(1-g)\sigma_s$ can be found.

Estimating optical parameters with a radiative transfer model would be valuable for the paper industry since it could increase the accuracy compared to using Kubelka–Munk. It could also resolve problems regarding ink characterization due to deficiencies of the Kubelka–Munk model for highly absorbing and optically thin samples [13]. Another important benefit would be the possibility of more accurate data exchange between the paper and printing industries. This is currently not possible, partly due to different instrument geometries, while the Kubelka–Munk parameters are geometry-dependent.

The parameter estimation problem consists in determining σ_s, σ_a and g (using the Henyey–Greenstein phase function (10.18)) from angle-resolved intensity measurements in chosen directions, $I(u_i, \varphi_i)$. Such measurements are available from special goniophotometers. The parameter estimation problem can thus be defined as

$$\textbf{find } \sigma_s, \sigma_a \text{ and } g \textbf{ given} \text{ the measurements } I(u_i, \varphi_i). \tag{10.22}$$

The parameter estimation problem is to find parameter values that minimize some distance measure between real measurements and model predictions. The

problem is usually over-determined with noisy measurements. The noise can be controlled by averaging over several measurements, but the parameter estimation is nevertheless a nonzero-residual problem.

One way to introduce the distance measure to minimize is through an objective function that sums squared errors [38], such as

$$F(x) = \frac{1}{2}||f(x)||^2 = \frac{1}{2}\sum_i f_i(x)^2 = \frac{1}{2}\sum_i \{M_i(x) - b_i\}^2, \tag{10.23}$$

where i denotes the respective quantity (e.g. intensity in a given direction), and x is the vector of parameters to be determined. Each element of the vector function $f(x)$ is the difference between model calculation $M_i(x)$ and measurement b_i. This formulation is statistically optimal if the measurement errors are normally distributed, which may reasonably be considered to be the case in a well-tuned industrial application. An obvious formulation of the parameter estimation problem is then $\min_x F(x)$, or the explicit least-squares formulation $\min_x \frac{1}{2}||f(x)||_2^2 = \min_x \frac{1}{2}\sum_i f_i(x)^2$. If one defines the set of permissible parameter combinations as

$$S = \{(\sigma_s, \sigma_a, g) : \sigma_s > 0, \sigma_a > 0, -1 < g < 1\}, \tag{10.24}$$

one can state the parameter estimation problem in various ways, which makes it possible to use different optimization methods to find a solution.

It should be pointed out that it is therefore often assumed that model output or sample measurements are given uniquely from medium parameters. This is often a reasonable assumption, for applied problems, and proof has been given for several reasonable cases [48], although not generally for the entire problem class. A quite general proof for the uniqueness of the solution to the inverse problem – given the existence of a solution to the forward problem – has been given by Choulli and Stefanov [49]. The uniqueness will also depend on the type of measurements done. The choice of experiment is of great importance, and the setup must be sensitive to the parameters to be estimated. There is still more work to be done to achieve full and general knowledge regarding existence and uniqueness for the general radiative transfer problem and its inverse.

Edström [38] performed a sensitivity analysis of the inverse problem with the objective function given by Eq. (10.23), and found expressions for the relative sensitivity of parameter x_i for change (perturbation) in measurement b_j. These show that even small perturbations in measurement j may result in large changes in parameter i. This may be the case for nonzero-residual problems with large curvature and for problems with a large residual, but it may also happen for zero-residual problems with ill-conditioned Jacobian or Hessian in the solution. Calculating material parameters from reflectance measurements is an outstanding issue in general radiative transfer problems. Finding a feasible starting point can in itself be a great problem in many applications.

Other types of sensitivity analysis – e.g. objective function surface plots, phase space plots and sensitivity matrices – give good insight into the character of the parameter estimation problem. Such studies will be valuable in the design of instruments for angle-resolved measurements and in analysis of the corresponding

measurement data. This is particularly important for the paper industry, where anisotropy is gradually becoming an issue. The sensitivity analysis points out which parameters are sensitive to what measurements. This shows the kind of accuracy that can be expected from the estimations, but it also helps to design new measurements with minimal impact of measurement noise. Numerical experiments can thus suggest better locations of measurement points to improve the final results of the parameter estimation. Feasibility and demands for industrial implementation of angle-resolved measurements are in initial discussions between European paper industries and institutes, and while online industrial implementation may lie a few years in the future, the use in research labs is already there and will necessarily increase to attack anisotropy issues.

The use of objective function surface plots is easily illustrated. The first pane of Fig. 10.8 indicates how cases with low opacity (defined as $Op = 100 \cdot R_0/R_\infty$) give well-conditioned problems with non-sensitive solutions for the d/0 problem (estimating only σ_s and σ_a from measurements of R_0 and R_∞). This is seen since the objective function surface is smooth and locally convex, and there is one distinct global minimum. Cases with high opacity (second pane of Fig. 10.8), on the other hand, seem to have an objective function surface that is flat in one or more directions, which shows that those cases give ill-conditioned problems with poor convergence (hard to find iteration steps in the optimization that give sufficient descent in the flat areas) or sensitive solutions (a small change in target value can give a large change in the parameter solution). Fig. 10.9 shows that the problem when estimating also the asymmetry factor g (now using angle-resolved intensity measurements) is not convex, and indicates that problems with higher g are more ill-conditioned. Although the objective function surface is smooth, the non-quadratic curvature and the local flatness along lines or curves indicate that the problem is ill-conditioned, possibly with poor convergence or sensitive solutions. It is also evident that there are ridges that will keep the optimization algorithm away from the area with the optimum, unless a sufficiently good starting point is provided. Numerical investigations show that the ill-conditioning is an even larger problem than indicated by Fig. 10.9, especially when noise is included.

10.3.2 Monte Carlo modeling

Monte Carlo (MC) models are becoming more frequently used in the same areas as radiative transfer models. They are inherently much slower and usually have a large number of parameters that are hard to determine (the parameter set also varies with the choice of physics to model), but faster computers make them more interesting. Although such models might not get fast enough for online use in a production environment, there are other advantages. Since it is possible to include any desired phenomena, one can avoid undesired approximations, and one can use physical rather than phenomenological parameters. This gives MC models a potentially great explanatory power.

In paper and printing applications, there are a number of important issues that cannot be easily handled with ordinary radiative transfer models, but that can be accommodated in an MC model. The obvious one is the possibility of modeling real structures, including the cellulose fiber network, filler particles, coating layers, colorants in inks etc., treating them as geometrical objects, and handling interactions

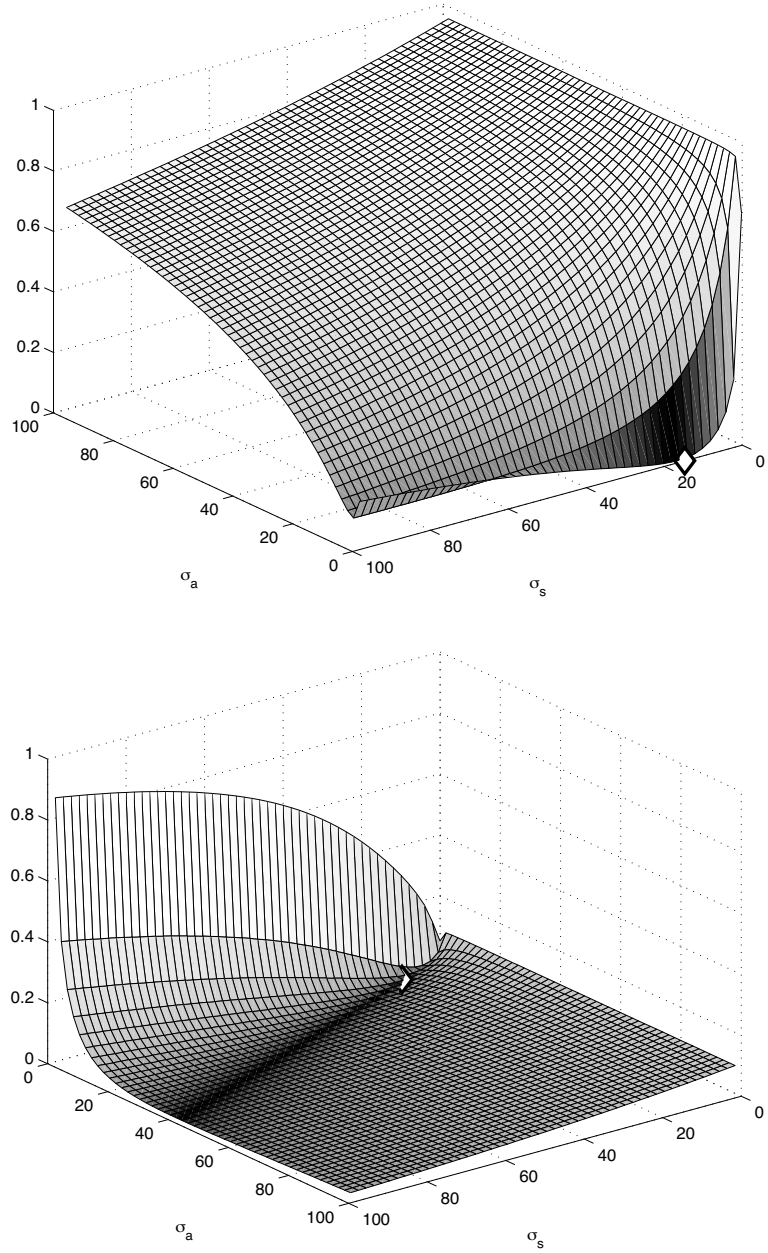


Fig. 10.8. Two test cases for the d/0 problem. Upper ($\sigma_s = 14.7$, $\sigma_a = 0.03$, $g = 0$, $w = 0.1$, opacity = 50.1%): The objective function surface is smooth and locally convex with one distinct local minimum. The problem should be well conditioned with a non-sensitive solution. The solution is, however, close to a boundary, which may give rise to problems. Lower ($\sigma_s = 14.0$, $\sigma_a = 5.6$, $g = 0$, $w = 0.1$, opacity = 95.5%): The objective function surface is smooth, but is locally flat along a line, possibly giving poor convergence or a sensitive solution. The diamonds indicate the points of convergence.

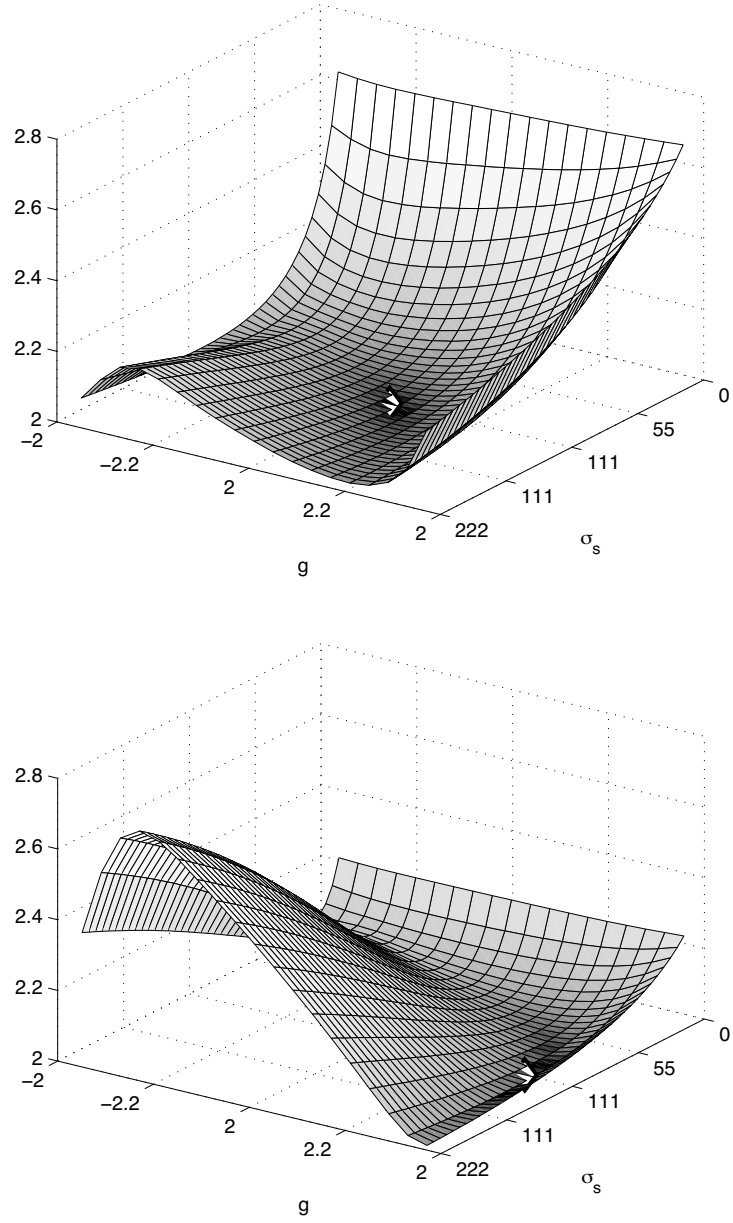


Fig. 10.9. Two test cases for the full problem using angle-resolved intensity measurements. Objective function surfaces as a function of σ_s and g with σ_a fixed. Upper: ($\sigma_s = 100$, $\sigma_a = 20$, $g = 0.05$, $w = 0.03$). Lower: ($\sigma_s = 100$, $\sigma_a = 20$, $g = 0.75$, $w = 0.03$). The objective function surface is smooth but non-convex, and is locally flat along a curve (especially for $g = 0.75$), possibly giving poor convergence or a sensitive solution. A poor starting point may lead to divergence because of the curvature. The diamonds indicate the points of convergence.

with fundamental physical laws. In MC models, a large number of light packages is sent to the sample according to a specified illumination, the light packages interact with the structure, and are finally collected in a specified detector. All kinds of statistics is then available for post-processing, and in principle there is no limit on what phenomena can be included in the simulation. One drawback of MC models are the long computation times, especially for highly scattering and optically thick media, but this may be acceptable for phenomena that are not possible to simulate in other ways. Another drawback is the use of many parameters, which need a large amount of measurement to determine. Yet some are not easily measured and need to be fitted, which does not always give unique or accurate results.

Among the most important issues that are almost exclusively simulated with MC models in the paper industry today are surface roughness and gloss, lateral scattering and optical dot gain, spatial variations and mottling, and fluorescence.

Most radiative transfer models do not consider surfaces explicitly, but handle only the bulk. When surfaces are included, it is most often a flat boundary. Rough surfaces are hard to model with differential models, which leaves only MC models. Gloss is an important feature of paper that comes from the roughness of every paper surface. The influence of gloss on visual appearance is large, so accurate modeling is important. Although several attempts have been made, there is still a lot to be done to fully understand and describe surface scattering and to distinguish it from bulk scattering in measurements.

Most radiative transfer models are not spatially resolved, or are not adapted to paper applications. Most printing is half-toned, meaning that images are achieved by printing a raster of small dots. Due to lateral scattering, the dots appear to grow, which is called optical dot gain, and this affects the perceived color. MC models are the only practical choice for accurate modeling of optical dot gain in the paper industry today.

A typical paper contains a lot of spatial variations due to agglomeration of paper constituents and uneven application of coatings and inks. This gives rise to optical mottling, gloss variations and other defects, which are directly coupled to the real physical structure of the paper. MC models are obviously the only tool for studies of such phenomena.

Fluorescence could in principle be handled also in a radiative transfer model. However, there is no suitable such model available for paper industry applications today. Therefore, MC models dominate here. However, radiative transfer models could well take over this niche.

An Open Source platform for Monte Carlo simulations of light scattering in paper is about to be launched in a joint Swedish project, where it will be free to contribute and share computer code and structure data. A web site will be released during 2009. Hopefully, the scientific community will find it valuable to use the tools for free, and will also hopefully contribute with further code development, addition of new components etc.

10.3.3 Impact on measurement systems and industry standards

Methods and standards for measuring and handling optical properties differ between the paper and graphic arts industries. This is due to the fact that development over the years has been in relative isolation between the industries, with different issues and goals in mind.

There is room for a lot of improvement and development in industrial application regarding measurement systems and standardization. If one wants to take advantage of new possibilities to model anisotropy and fluorescence, it is going to be necessary to agree upon and standardize goniophotometric and bispectral measurement devices. But already for existing devices, calibration routines, UV content of illumination, area of measurement contra illumination and so on need to be better specified in order to facilitate accurate data exchange. There is a need for overall harmonization of standards since there are so many that most applied users do not even know which standards do apply in a given setting, and there are also gaps, overlaps and inconsistencies.

The wish to model the real structure of paper and printed paper products instead of homogenized approximations will drive the development of MC models with real explanative power, which in turn will give an increased need for characterization methods.

10.4 Discussion

On the more theoretical side, it can be noted that not much is known about the uniqueness of solutions to the inverse radiative transfer problem (practically, the uniqueness clearly depends on the chosen measurement setup, and the experiment must be optimized with respect to the parameters to be retrieved from indirect measurements). A functional analytic approach would find a rich research area here. Since the use of models of this type is growing rapidly in various applied sciences, studies of this kind are called for.

On the more applied side is the investigation and utilization of nanomaterials. This follows the general trend in other sciences. Since nanomaterials can have extreme properties, they can probably find use in both surface treatment and in inks. One recent example is the discovery that ordinary candle soot contains multicolor fluorescent carbon nanoparticles [50], and it would be interesting to see what use this could have for paper whiteness or special inks. Since the size of nanoparticles is of the same order as the wavelength of visible light, the interaction of light with these particles may need different modeling strategies than the ones used in the area today. With the extreme properties of new nanomaterials follows the challenge to characterize them. Characterization will be needed both optically (scattering and absorption parameters, asymmetry factor, fluorescence properties) and otherwise (position in paper structure, effect on other properties such as chemistry, etc.).

Spatially resolved radiative transfer models and corresponding forward and inverse solution algorithms are interesting within both mathematics and applied sciences. Theoretically, even less is known here, and applications need more powerful tools than the rather crude finite element methods in use today. Various imaging,

localizing, characterizing and optimizing problems in most areas where radiative transfer is used today would greatly benefit from a breakthrough here. In printing applications, the possibility of correctly modeling optical dot gain has long been waited for.

Monte Carlo models are becoming more frequently used in the same areas as radiative transfer models. Although they are much slower and often have a large number of parameters that are not easily determined, the development of faster computers will make them increasingly interesting. Furthermore, the possibility of including any desired phenomena, to avoid undesired approximations and to use physical rather than phenomenological parameters, gives Monte Carlo models potentially great explanatory power. There are ongoing activities in this area. An outstanding problem then is parameter estimation in Monte Carlo models. Their genuinely discrete nature and their long computation times are an obstacle to their full use in applied sciences.

Looking more specifically at the paper optics area, it can be noted that there are at least two key areas that need more modeling development. One is fluorescence phenomena, which have a large impact on visual appearance and color reproduction. Yet, the treatment of fluorescence in standards, practice and models is varying, poor or non-existent. The UV spectrum of incident light in measurement devices needs to be standardized, and bispectral measurement devices need to be standardized and widely used for fluorescence characterization. The other area is surface roughness and gloss, which also have a large impact on visual appearance. Although several attempts have been made to model this, a lot remains to be done to fully understand and describe surface scattering and to distinguish it from bulk scattering in measurements.

Very little is written and known about numerical values of phase function moments χ_l or the asymmetry factor g (or the reduced scattering coefficient $(1-g)\sigma_s$) for materials like paper, and the question of anisotropy and the utilization of the asymmetry factor is new to the paper industry. Unlike several radiative transfer applications, NIR/IR is not relevant, only the visible range (and UV via fluorescence) is of importance. Spectral goniophotometer measurements of various paper samples are needed for a broad evaluation of the new parameter estimation methods for the asymmetry factor, but also of the spectral dependence and the numerical values of the asymmetry factor. The inverse problem should be studied for examining the influence of measurement errors or noise on the retrieved parameters; this would make it possible to design experiments and measurement setups with minimal influence from measurement errors. The possibility of using the asymmetry factor as a design variable opens for new opportunities in the design of new paper products.

Most industrial models of today use phenomenological and model dependent parameters, and this has some disadvantages. Since the parameters are not physically objective, they are not necessarily exchangeable between situations. Measuring the same samples with different instruments may give not only different values but also different ranking, which is not satisfactory. One example of this is estimation of optical properties from measurements using d/0 and 45/0 geometry. Also, phenomena that might be considered physically linear do not necessarily appear so since the phenomenological parameters are not the actual physical quantities. This also diminishes the explanatory power of the models, since the parameters

are actually defined by the models and therefore often have no absolute and objective meaning outside. It is therefore desirable to use physically objective material parameters, real paper properties and real process parameters. This makes data more exchangeable since they are independent of model and measurement geometry. This also gives the models more explanatory power, as they convey more of the physical reality. It is a clear trend that any new optical models introduced will need to use physically objective material parameters instead of phenomenological model parameters.

10.5 Conclusions

How the paper and printing industries successfully use the Kubelka–Munk model for interpretation of reflectance measurements has been discussed. Simulation and modeling is used for prediction in the development of new products, for quality control in production, and for data exchange in trading situations.

It has been noted that there is room for improvement regarding accuracy and efficiency in today's use, and that new and higher demands on knowledge, accuracy and range of applicability drive the need for new simulation models. Suitable models and measurements are in principle available, but relatively little work of the scientific community is directly adapted to the specific industrial needs and demands.

Examples of current challenges and future work have been outlined, and it is evident that much progress can be made to advance the understanding of several industrially relevant phenomena through proper application of modeling and measurements.

References

1. ISO 2469, *Paper, board and pulps – Measurement of diffuse reflectance factor*, International Organization for Standardization, Geneva, Switzerland (1994).
2. ISO 2470, *Paper, board and pulps – Measurement of diffuse blue reflectance factor (ISO brightness)*, International Organization for Standardization, Geneva, Switzerland (1999).
3. ISO 2471, *Paper and board – Determination of opacity (paper backing) – Diffuse reflectance method*, International Organization for Standardization, Geneva, Switzerland (1998).
4. ISO 9416, *Paper – Determination of light scattering and absorption coefficients (using Kubelka–Munk theory)*, International Organization for Standardization, Geneva, Switzerland (1998).
5. Kubelka, P. and F. Munk, 1931: Ein Beitrag zur Optik der Farbanstriche, *Z. Tech. Phys.*, **11a**, 593–601.
6. Kubelka, P., 1948: New contributions to the optics of intensely light-scattering materials. Part I, *J. Opt. Soc. Amer.*, **38**, 448–457.
7. Kubelka, P., 1954: New contributions to the optics of intensely light-scattering materials. Part II, *J. Opt. Soc. Amer.*, **44**, 330–335.
8. van den Akker, J. A., 1949: Scattering and absorption of light in paper and other diffusing media, *TAPPI*, **32**, 498–501.

9. Foote, W. J., 1939: An investigation of the fundamental scattering and absorption coefficients of dyed handsheets, *Paper Trade Journal*, **109**, 397–404.
10. Nordman, L., P. Aaltonen and T. Makkonen, 1966: Relationship between mechanical and optical properties of paper affected by web consolidation, *Trans. Symp. 'Consolidation of the Paper Web'*, vol. 2, 909–927.
11. Moldenius, S., 1983: Light absorption coefficient spectra of hydrogen peroxide bleached mechanical pulp, *Paperi Puu*, **65**, 747–756.
12. Rundlöf, M. and J. A. Bristow, 1997: A note concerning the interaction between light scattering and light absorption in the application of the Kubelka–Munk equations, *J. Pulp Paper Sci.*, **23**, 220–223.
13. Neuman, M., 2005: *Anisotropic Reflectance from Paper – Measurements, Simulations and Analysis*, Master's thesis, Umeå University, Sweden.
14. van den Akker, J. A., 1968: Theory of some of the discrepancies observed in application of the Kubelka–Munk equations to particulate systems, *Modern Aspects of Reflectance Spectroscopy*, W. W. Wendlandt, Ed., Plenum Press, New York, 27–46.
15. Koukoulas, A. A. and B. D. Jordan, 1997: Effect of strong absorption on the Kubelka–Munk scattering coefficient, *J. Pulp Paper Sci.*, **23**, 224–232.
16. van den Akker, J. A., 1966: Discussion on 'Relationships between mechanical and optical properties of paper affected by web consolidation', *Trans. Symp. 'Consolidation of the Paper Web'*, vol. 2 948–950.
17. Granberg, H. and P. Edström, 2003: Quantification of the intrinsic error of the Kubelka–Munk model caused by strong light absorption, *J. Pulp Paper Sci.*, **29**, 386–390.
18. Nobbs, J. H., 1985: Kubelka–Munk theory and the prediction of reflectance, *Rev. Prog. Coloration*, **15**, 66–75.
19. Granberg, H. and M.-C. Béland, 2004: Modelling the angle-dependent light scattering from sheets of pulp fibre fragments, *Nordic Pulp Paper Res. J.*, **19**, 354–359.
20. Edström, P., 2005: A fast and stable solution method for the radiative transfer problem, *SIAM Rev.*, **47**, 447–468.
21. Kokhanovsky, A. A., 2006: Asymptotic radiative transfer, *Light Scattering Reviews*, vol. 1, A. A. Kokhanovsky, Ed., Springer, Berlin, 253–289.
22. Kokhanovsky, A. A. and L. G. Sokoletsky, 2006: Reflection of light from semi-infinite absorbing turbid media. Part 1: Spherical albedo, *Color Res. Appl.*, **31**, 491–497.
23. Kokhanovsky, A. A. and L. G. Sokoletsky, 2006: Reflection of light from semi-infinite absorbing turbid media. Part 2: Plane albedo and reflection function, *Color Res. Appl.*, **31**, 498–509.
24. Philips-Invernizzi, B., D. Dupont and C. CazÈ, 2001: Bibliographical review for reflectance of diffusing media, *Opt. Eng.*, **40**, 1082–1092.
25. Murphy, A. B., 2007: Optical properties of an optically rough coating from inversion of diffuse reflectance measurements, *Appl. Opt.*, **46**, 3133–3143.
26. Kokhanovsky, A. A., 2007: Physical interpretation and accuracy of the Kubelka–Munk theory, *J. Phys. D: Appl. Phys.*, **40**, 2210–2216.
27. Kokhanovsky, A. A. and I. Hopkinson, 2008: Some analytical approximations to radiative transfer theory and their application for the analysis of reflectance data, *J. Opt. A: Pure Appl. Opt.*, **10**, 035001.
28. Mudgett, P. S. and L. W. Richards, 1971: Multiple scattering calculations for technology, *Appl. Opt.*, **10**, 1485–1502.
29. Mudgett, P. S. and L. W. Richards, 1972: Multiple scattering calculations for technology II, *J. Colloid Interf. Sci.*, **39**, 551–567.
30. van de Hulst, H. C., 1980: *Multiple Light Scattering. Tables, Formulas and Applications.* vol. 2, Academic Press, New York.

31. Yang, L. and B. Kruse, 2004: Revised Kubelka–Munk theory. I. Theory and applications, *J. Opt. Soc. Am. A*, **21**, 1933–1941.
32. Yang, L., B. Kruse and S. J. Miklavcic, 2004: Revised Kubelka–Munk theory. II. Unified framework for homogenous and inhomogenous optical media, *J. Opt. Soc. Am. A*, **21**, 1942–1952.
33. Yang, L. and S. J. Miklavcic, 2005: A theory of light propagation incorporating scattering and absorption in turbid media, *Opt. Lett.*, **30**, 792–794.
34. Yang, L. and S. J. Miklavcic, 2005: Revised Kubelka–Munk theory. III. A general theory of light propagation in scattering and absorptive media, *J. Opt. Soc. Am. A*, **22**, 1866–1873.
35. Edström, P., 2007: Examination of the revised Kubelka–Munk theory: considerations of modeling strategies, *J. Opt. Soc. Am. A*, **24**, 548–556.
36. Kokhanovsky, A. A., 2002: Statistical properties of a photon gas in random media, *Phys. Rev. E*, **66**, 037601.
37. Edström, P., 2009: Numerical performance of stability enhancing and speed increasing steps in radiative transfer solution methods, *J. Comput. Appl. Math.*, **228**, 104–114.
38. Edström, P., 2008: A two-phase parameter estimation method for radiative transfer problems in paper industry applications, *Inverse Probl. Sci. Eng.*, **16**, 927–951.
39. Feng, T., P. Edström and M. Gulliksson, 2007: Levenberg–Marquardt methods for parameter estimation problems in the radiative transfer equation, *Inverse Problems*, **23**, 879–891.
40. Edström, P., 2004: Comparison of the DORT2002 radiative transfer solution method and the Kubelka–Munk model, *Nordic Pulp Paper Res. J.*, **19**, 397–403.
41. Lord Rayleigh, 1871: On the light from the sky, its polarization and colour, *Philos. Mag.*, **41**, 107–120, 274–279. (Reprinted in *Scientific Papers by Lord Rayleigh*, vol. I: 1869–1881, No. 8, Dover, New York, 1964.)
42. Mie, G., 1908: Beiträge zur Optik trüber Medien, Speziell Kolloidaler Metallösungen, *Ann. Phys.*, **25**, 377–445.
43. Henyey, L. G. and J. L. Greenstein, 1941: Diffuse radiation in the galaxy, *Astrophys. J.*, **93**, 70–83.
44. Chandrasekhar, S., 1960: *Radiative Transfer*, Dover, New York.
45. van Gemert, M. J. C. and W. M. Star, 1987: Relations between the Kubelka–Munk and the transport equation models for anisotropic scattering, *Lasers Life Sci.*, **1**, 287–298.
46. Prahl, S. A., M. J. C. van Gemert and A. J. Welch, 1993: Determining the optical properties of turbid media using the adding-doubling method, *Appl. Opt.*, **32**, 559–568.
47. Joshi, N., C. Donner and H. W. Jensen, 2006: Noninvasive measurement of scattering anisotropy in turbid materials by nonnormal incident illumination, *Opt. Lett.*, **31**, 936–938.
48. Case, K. M. and P. F. Zweifel, 1967: *Linear Transport Theory*, Addison-Wesley, Reading, MA.
49. Choulli, M. and P. Stefanov, 1998: An inverse boundary value problem for the stationary transport eEquation, *Osaka J. Math*, **36**, 87–104.
50. Liu, H., T. Ye and C. Mao, 2007: Fluorescent carbon nanoparticles derived from candle soot, *Angew. Chem. Int. Ed.*, **46**, 6473–6475.

11 Coherent backscattering in planetary regoliths

Karri Muinonen, Jani Tyynelä, Evgenij Zubko, and Gorden Videen

11.1 Introduction

Atmosphereless solar-system objects exhibit two ubiquitous light-scattering phenomena at small solar phase angles (sun–object–observer angle α): first, the opposition effect in the intensity of scattered sunlight (e.g., [1]); and, second, the negative degree of linear polarization $(I_\perp - I_\parallel)/(I_\perp + I_\parallel)$. Here $I_\parallel$ denotes the intensity component parallel to the scattering plane defined by the Sun, the object, and the observer and $I_\perp$ denotes the component perpendicular to that plane [2].

Narrow opposition effects for the bright E-class asteroids (44) Nysa and (64) Angelina have been discovered by Harris et al. [3]; whereas, sharp polarimetric phase effects have been discovered for the same objects by Rosenbush et al. [4, 5]. Observations of transneptunian objects made at the Very Large Telescope of the European Southern Observatory have yielded both sharp and shallow polarimetric phase effects at phase angles within 1.5 degrees of exact opposition (e.g., [6–8]). A peculiar asymmetric polarization curve with a large phase angle of polarization minimum was discovered for asteroid (234) Barbara by Cellino et al. [9]. For reviews on phase effects observed for solar-system objects, see [10–12]. Observations on the photometric and polarimetric phase effects constitute often the only observational data available on a given solar-system object, so it is important to interpret these data correctly.

The physical interpretation of these observations requires an understanding of the intricate interplay of single scattering and multiple scattering in planetary regoliths. The single scatterers can sometimes be identified to be the small particles in the objects; whereas, sometimes, the scatterers are the inhomogeneities within otherwise homogeneous particles that are themselves substantially larger than the wavelength. Multiple scattering among the single scatterers can result in smoothing of the scattering characteristics for some objects but sharpening of the characteristics for others.

There are no exact electromagnetic solutions for light scattering by what the surfaces of atmosphereless solar-system objects are presumably composed of, i.e., close-packed random media of inhomogeneous particles large compared to the wavelength. The direct problem of computing scattering characteristics for media with well specified physical properties poses a major challenge. In the inverse problem

of deriving information about the physical properties of planetary regoliths based on the observational data available, it has been mandatory to make use of simplified scattering models. For example, the popular photometric models (e.g., [13–15]) for the inverse problem account for the single-particle albedo and phase function, the volume density (or fraction) of particles, and the roughness of the interface between the scattering medium and the free space. It is widely accepted that, analyzed separately, disk-integrated photometric or polarimetric phase effects provide two ill-posed inverse problems. Analyzing the two data sets simultaneously and/or moving to disk-resolved analyses can at least partially remove the ambiguities.

Particle morphology is one primary factor in determining how incident light is scattered and absorbed by particles. Conversely, knowledge of the differences and similarities of particle extinction, scattering, and absorption cross-sections, asymmetry parameters, and scattering phase matrices may provide clues to the morphology and is of key importance in remote-sensing studies of small solar-system objects such as asteroids and comets.

As a summary for physical modeling, multiple scattering in planetary regoliths depends on five groups of physical parameters: first, surface roughness in length scales of several wavelengths and large numbers of particles cause shadowing effects; second, the porosity of the particulate medium causes shadowing effects; third, particle size introduces its signature into the scattering characteristics; fourth, particle shape plays an important role in determining the detailed structure of the particle scattering matrix; and, fifth, the optical properties of the material of which the particles are composed (including inhomogeneous media) dictate the material response to incident light, thereby also influencing the scattering matrix.

Substantial advances in theories for the physical causes of opposition phenomena have been made in the past 25 years: the coherent-backscattering mechanism (CBM) has been established to contribute to the photometry and polarimetry at small phase angles [16–20], challenging the traditional shadowing-mechanism (SM) interpretations. Whereas CBM is a multiple-scattering mechanism for scattering orders higher than the first, SM is a first-order multiple-scattering mechanism. Both CBM and SM contribute to the opposition effects observed for atmosphereless objects. Whereas CBM can contribute to the negative polarization branch, it continues to be a matter of some controversy whether SM, as presented in [21], can affect the polarization branch. There are several reviews of the mechanisms for backscattering phenomena of atmosphereless solar system bodies (e.g., [22–24]), and we refer the reader to these reviews for a more detailed history of, e.g., CBM and SM.

Coherent backscattering by plane-parallel media of spherical scatterers has been treated in [25] using numerical Monte Carlo methods, accounting for polarization. Based on that treatment, methods have been developed for spherical media of Rayleigh scatterers to interpret the polarimetric and photometric observations obtained for transneptunian objects (e.g., [6, 10]). These methods are described in detail in the present chapter.

In planetary regoliths, SM for opposition brightening is relevant for essentially all length scales much larger than the wavelength of incident light. We can distinguish between two main geometric factors affecting SM: first, the rough interface between the regolith and the free space; and, second, the internal porous structure of

the regolith. In both cases, the opposition brightening is due to the fact that a ray of light penetrating into the scattering medium and incident on a certain particle can always emerge back along the path of incidence; whereas, in other directions, the emerging ray can be blocked by other particles. Porous-volume shadowing depends mainly on the volume density of the scattering medium; whereas, rough-surface shadowing depends mainly on surface roughness parameters. For recent work on rough-surface shadowing for fractional-Brownian-motion surfaces and on combined rough-surface and porous-volume shadowing for closely packed random media of spheres with varying sizes, see, e.g., [26] and [27], respectively. For media involving Gaussian random surfaces with Gaussian correlation functions, see, e.g., [28] and [29].

Wide backscattering peaks and negative polarization branches have been detected consistently in numerical simulations of light scattering by irregular wavelength-scale particles (e.g., [30–32]). The phenomena are present for compact irregular particles as well as for irregular aggregates of constituent spherical or nonspherical particles. Recently, we have succeeded in uncovering internal-field characteristics that give rise to such polarization and intensity signatures [32–36] and thus have introduced what we call a single-scattering polarization and intensity mechanism [37]. Note that with the single-scattering mechanism, it has not been our central goal to explain the detailed characteristics of the internal fields themselves but rather concentrate on how the internal fields map to the scattered fields. As to the light-scattering experiments, by measuring the single-particle scattering characteristics and those of a close-packed particulate medium of similar single particles, it has been established that the particulate media continue to exhibit single-particle polarization characteristics but that these characteristics are neutralized (e.g., [38]).

In section 11.2, we describe the relevant scattering theory, discuss example scattering phase matrices for Gaussian-random-sphere particles (G-sphere particles) and agglomerated debris particles, analyze internal-field characteristics for spherical and G-sphere particles, and outline interference processes capable of explaining the intensity and polarization characteristics at intermediate scattering angles as well as close to backward scattering. We close section 11.2 by parameterizing the amplitude scattering and scattering phase matrices so as to reproduce rapidly realistic matrices for further use in multiple-scattering computations. In section 11.3, we review the coherent-backscattering mechanism for enhanced backscattering and for negative polarization near backscattering. For applications not concerning polarization, we present a scalar approximation in coherent backscattering by spherical random media of diametrical optical thickness approaching infinity, analyzing in detail various implications of varying single-scattering albedo and phase function on the characteristics of enhanced backscattering. We continue by utilizing the numerical methods for coherent backscattering, including polarization, for a study of backscattering characteristics of spherical random media composed of Rayleigh scatterers. In section 11.4, we describe a physical model for the polarimetric observations of solar-system objects. We construct the model on the basis of the single-scattering parameterization to be given in section 11.2, including averaging over scatterer size via numerical integration. In order to substantiate the applicability of the physical model, we provide an example least-squares fit to the

polarimetric observations of asteroid (1) Ceres. We show computations of coherent backscattering by spherical random media of scatterers utilizing the parameterized amplitude scattering and scattering phase matrices, thus paving the way for simultaneous physical modeling of photometric and polarimetric observations. We close the chapter with conclusions in section 11.5.

11.2 Single-particle light scattering

11.2.1 Scattering matrix, cross-section, and asymmetry parameters

Consider scattering of light by a single particle in free space. The scattering characteristics depend on the particle size, shape, complex refractive index, and orientation. In what follows, we study the average scattering characteristics resulting from a large number of single-particle scattering problems and, in the limit of an infinite number of problems, there are equal numbers of particles and their mirror particles involved in ensemble averaging, both in random orientation.

The scattering phase matrix $\boldsymbol{P} = \boldsymbol{P}(\theta)$ ($\theta = \pi - \alpha$ is the scattering angle) relates the Stokes vectors of the incident and scattered light ($\boldsymbol{I}_{\rm inc}$ and $\boldsymbol{I}_{\rm sca}$, respectively),

$$\begin{aligned} \boldsymbol{I}_{\rm sca} &= \frac{\sigma_{\rm sca}}{4\pi R^2}\,\boldsymbol{P}\cdot\boldsymbol{I}_{\rm inc}, \\ N_{11} &= \int_{4\pi}\frac{d\Omega}{4\pi}\,P_{11}(\theta) = \frac{1}{2}\int_0^\pi d\theta\,\sin\theta P_{11}(\theta) = 1, \end{aligned} \tag{11.1}$$

where $\sigma_{\rm sca}$ and P_{11} are the ensemble-averaged scattering cross-section and phase function, $N_{11} = 1$ is the norm of P_{11}, and R is the distance between the scatterer and the observer. The scattering phase matrix is block-diagonal with P_{11}, $P_{12} = P_{21}$, P_{22}, P_{33}, $P_{34} = -P_{43}$, and P_{44} as its non-vanishing elements. Ensemble-averaging is carried out for cross-section-weighted elements, e.g., $P_{11} = \langle\sigma_{\rm sca}^{(1)}P_{11}^{(1)}\rangle/\sigma_{\rm sca}$, where the superscript '(1)' refers to an individual sample particle. In principle, ensemble averaging entails an infinite number of sample particles but, in practice, a large but finite number of particles suffices.

The degree of linear polarization P for unpolarized incident light and the depolarization ratio D are defined as

$$P(\theta) = -\frac{P_{12}(\theta)}{P_{11}(\theta)},\quad D(\theta) = 1 - \frac{P_{22}(\theta)}{P_{11}(\theta)}. \tag{11.2}$$

The polarization norm is [25]

$$N_{12} = -\int_{4\pi}\frac{d\Omega}{4\pi}\,P_{12}(\theta) = -\frac{1}{2}\int_0^\pi d\theta\,\sin\theta P_{12}(\theta), \tag{11.3}$$

which indicates, first, whether the scatterers are net positively or negatively polarizing and, second, how efficiently they polarize incident unpolarized light. For Rayleigh scatterers, $N_{12} = \frac{1}{2}$ (cf. the phase function norm $N_{11} = 1$).

The phase-function and polarization asymmetry parameters are defined as [25]

$$g_{11} = \frac{1}{2}\int_0^\pi d\theta \, \sin\theta \, \cos\theta P_{11}(\theta),$$
$$g_{12} = -\frac{1}{2}\int_0^\pi d\theta \, \sin\theta \, \cos\theta P_{12}(\theta). \quad (11.4)$$

For Rayleigh scatterers, $g_{11} = g_{12} = 0$. Together, N_{12} and g_{12} indicate the net positive or negative polarization characteristics in the forward and backward scattering regimes.

11.2.2 Scattering by Gaussian-random-sphere and agglomerated-debris particles

The discrete-dipole approximation (DDA) is a flexible tool that can be used to find the numerical solution of scattering by irregularly shaped particles (e.g., [39–43]). We show results calculated using the DDA code written by Zubko et al. [43]. Among the strengths of DDA is its conceptual clarity that allows in-depth studies of the underlying physical mechanisms (cf. ray-optics approximation for particles large compared to the wavelength).

Regularly shaped particles are exceptional in nature. The G-sphere particle allows the modeling of compact irregular particle shapes with a small number of statistical parameters [44–46]. In the present context, there are two such parameters: the relative standard deviation of the radial distance (σ) and the power-law index of the covariance function (ν) for logarithmic radial distances. The stochastic geometry of the agglomerated debris particle [43] allows for the modeling of the structures for porous particles. For sample G-sphere and agglomerated-debris particles, see Fig. 11.1.

Light scattering by G-sphere particles has been studied using various approximations, that is, the ray-optics [45–51], Rayleigh-volume and Rayleigh–Gans [52], Rayleigh-ellipsoid [53], and second-order perturbation-series approximations [44,

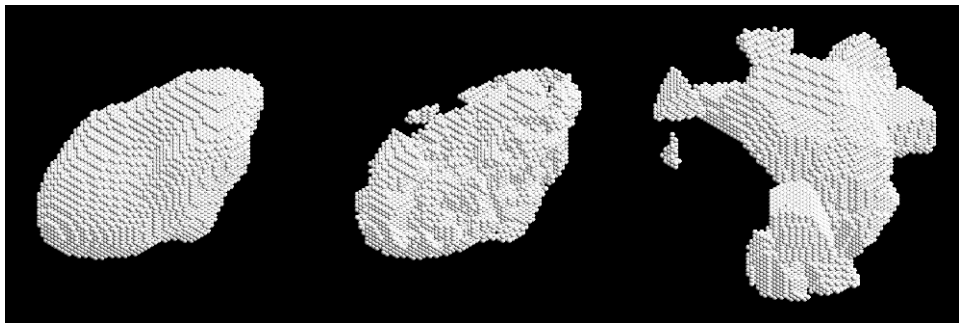


Fig. 11.1. Samples of Gaussian-random-sphere (G-sphere, left), roughened G-sphere (middle), and agglomerated-debris particles (right) all discretized for numerical computation using the discrete-dipole approximation.

54]. As to treatments that are close to being exact, scattering by G-sphere particles has been studied using the volume-integral-equation and DDA methods [30, 32, 35, 55, 56], and using the finite-difference time-domain method [57].

Muinonen et al. [32] and Zubko et al. [35] studied light scattering from G-sphere particles having circumscribing-sphere size parameters $x_{\rm cs} = 2$, 4, 6, 8, 10, 12, and 14, refractive indices $m = 1.6 + i0.0005$, and 1.313 using the DDA. The size parameters $x_{\rm cs}$ correspond to more commonly used size parameter $x = ka \in [1, 7]$, where the size of the G-sphere particle is defined by the mean radial distance a. It is important to notice that the sample particles drawn from the ensemble with given a typically vary by their volume, yielding a distribution of equal-volume-sphere size parameters. Two complex refractive indices (m) are treated in the simulations, that of pure water ice $m = 1.313$ and that of silicates $m = 1.6 + i0.0005$. The relative standard deviations are $\sigma = 0.05$, 0.1, 0.15, 0.2, and 0.245. The power-law index of the covariance function is $\nu = 4$.

These recent studies of light scattering by G-sphere particles have been accompanied by a closely related study: light scattering by harmonic G-sphere particles has been examined in [56] in order to map the surface-roughness effects for wavelength-scale particles. In [35], the surface-roughness effects have been studied for G-sphere particles with their surface layers additionally roughened in varying angular and radial scales. The two studies complement one another. Harmonic G-sphere particles have allowed a systematic study of the intrinsic roughness of the particles in precisely defined scales; whereas, introducing additional roughness on the irregular particles has removed the limitation of sphericity of the original particle.

In Fig. 11.2, we show the minimum polarization $P_{\rm min}$ and maximum polarization $P_{\rm max}$ for ensemble-averaged G-sphere, roughened G-sphere, and agglomerated-debris particles as a function of $x_{\rm cs}$. $P_{\rm min}$ and $P_{\rm max}$ for the agglomerated-debris particles lie in between the two G-sphere-particle values. In Fig. 11.3, we show the scattering phase functions and degrees of linear polarization for a subset of the aforedescribed particles. There are striking similarities in all of the angular dependences shown: all of the scattering phase functions show increased backscattering, the roughened G-sphere particle showing the smoothest dependence. The G-sphere cases show more pronounced negative polarization at intermediate scattering angles for both refractive indices.

11.2.3 Internal vs. scattered fields

In order to explain the intensity and polarization characteristics seen for G-sphere and agglomerated-debris particles shown, for example, in Fig. 11.3, we have carried out studies of the interrelations between internal and scattered fields for miscellaneous scatterers. Along the way, we have found that the internal fields of scatterers exhibit both intriguingly regular and irregular structures. Whereas the mapping of the internal field to the scattered field is well known from DDA simulations, the interference phenomena involved turn out to vary in complexity and character. Among the most common characteristics resulting from internal-field structure is the well known dependence of extinction cross-section of small particles on the central phase difference $\rho = |m - 1|kD$ (e.g., [58]), where m is the complex refractive

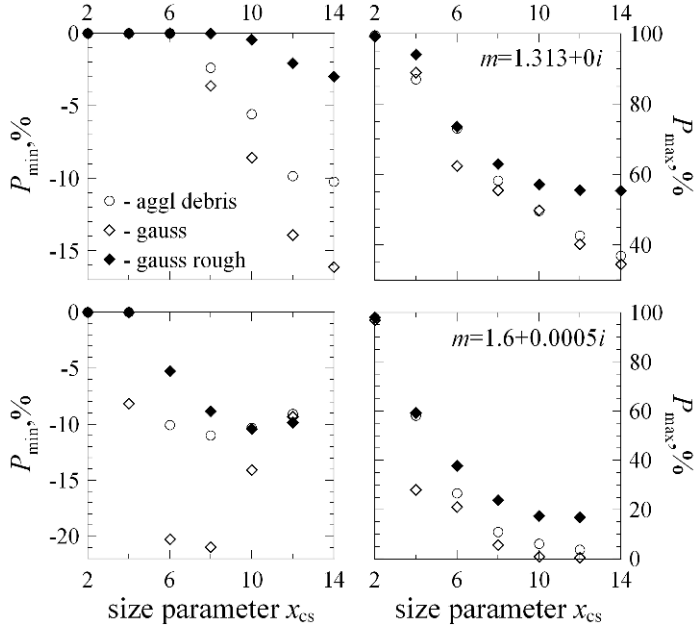


Fig. 11.2. The minimum polarization $P_{\min}$ and maximum polarization $P_{\max}$ for ensemble-averaged G-sphere, roughened G-sphere, and agglomerated-debris particles as a function of the circumscribing-sphere size parameter.

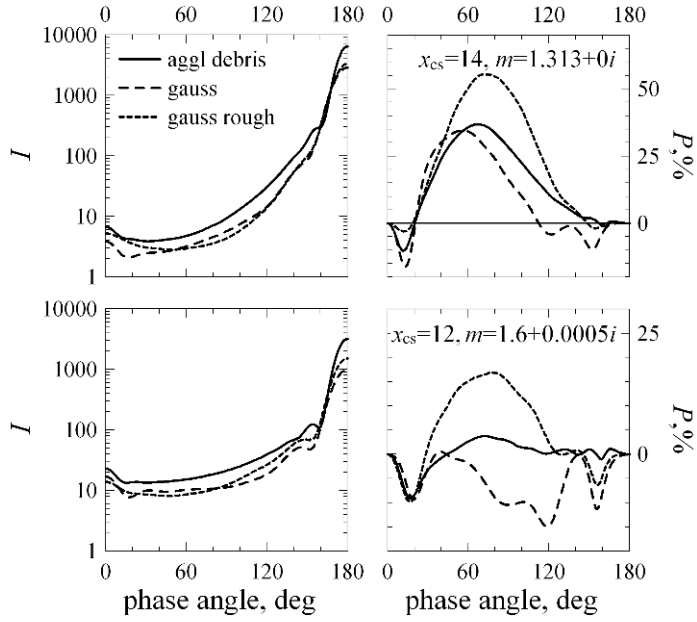


Fig. 11.3. The scattering phase functions I and degrees of linear polarization P for ensemble-averaged G-sphere, roughened G-sphere, and agglomerated-debris particles as a function of the circumscribing-sphere size parameter and refractive index.

index of the scattering particle, k is the wave number in free space, and D is the distance measured through the center of the particle.

The internal fields are readily available from DDA computations for the discrete dipoles involved, allowing for detailed studies of their contributions to the scattered fields. We point out that the following studies for spherical, G-sphere, and agglomerated-debris particles have been motivated by the phase-randomization study carried out by Zubko et al. [59] (see subsection 11.2.4).

Tyynelä et al. [33, 34] have studied the internal-field characteristics and interference effects of spherical and G-sphere particles. Each individual particle has been approximated with a cubic array of electric dipoles for which the electric field values have been adopted either from the exact Mie solution for spherical particles or the DDA solution for G-sphere particles. The scattering plane has been taken to be the xz-plane. In order to obtain the full scattering characteristics, the internal fields have been resolved for two perpendicular linear polarization states of the incoming wave; that is, the incident wave has been polarized linearly either along the x-axis or the y-axis, either parallel or perpendicular to the scattering plane.

The internal fields of spherical particles exhibit a lateral structure in agreement with the central phase difference ρ described above. Towards the perimeter of the projection of the sphere into the plane perpendicular to the direction of the incoming wave, the lateral structure gradually becomes curved so that, on the boundary, it approximately matches the wavelength of the incident plane wave in free space. The mapping of these internal fields to the scattered fields is carried out with a regular grating with the wavelength being that in free space. The differing wave structures of the internal fields and the mapping to the scattered field is the source of the challenges in constraining the causes for the intensity and polarization phenomena observed. Note, first, that the analysis has not been so much concerned with the question of why the internal fields are what they are; instead, the focus has been on understanding the main relevant interference effects involved. Note, second, that the internal fields exhibit structures mimicking those from refraction in geometric optics, even though geometric optics is strictly not valid for wavelength-scale particles.

For spherical particles, Tyynelä et al. [33] have continued by dividing the particle interior into quadrants along planes that are parallel to the propagation direction of the incident wave. They have further divided each quadrant into four slices defined by taking a regular square lattice through the quadrant again in the direction of propagation of the incident wave. They have then gradually neglected the interference effects between the different parts. They have considered size parameters $x = 2$, 4, 6, and 8, and refractive indices $m = 1.55$, $1.55 + \mathrm{i}0.01$, 1.33, $1.33 + \mathrm{i}0.01$, $1/1.55$, $1/1.33$, and $2.0 + \mathrm{i}2.0$.

For $m = 1.55$, the longitudinal component of the internal fields has two predominating maxima in the forward part of the particle that are separated by a distance nearly half the free-space wavelength, and also smaller localized maxima, which become more numerous as size increases. For a y-polarized incident wave, the two maxima produce a null effect in the scattering plane due to destructive interference. For an x-polarized incident wave, the two maxima interfere constructively for $\theta = 90^\circ$, providing a contribution to both intensity and polarization at intermediate scattering angles. For $m = 1/1.55$, the internal field is concentrated

on the backward part of the particle, and is weaker than for $m = 1.55$ both in strength and structure.

When the longitudinal internal-field component is omitted, the degree of linear polarization becomes positive for all scattering angles except for the angles near the backscattering direction. This can be seen in Fig. 11.4 (right panel) for $x = 4$ and $m = 1.55$. When the interior is further divided into incoherently radiating quadrants, polarization becomes positive for all scattering angles except for large size parameters ($x > 4$). When the interior is divided into sixteen parts, the po-

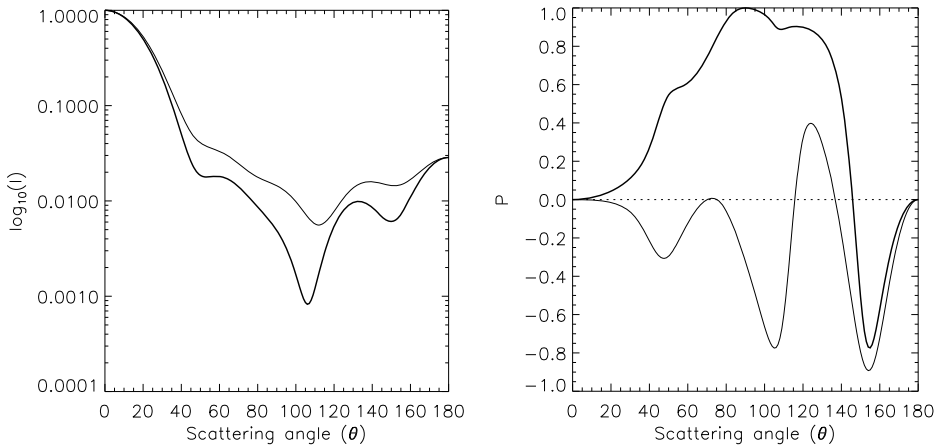


Fig. 11.4. Intensity $I_{\parallel}+I_{\perp}$ (left) and degree of linear polarization $P = (I_{\perp}-I_{\parallel})/(I_{\perp}+I_{\parallel})$ (right) in the case of the unmodified internal field (thin solid line) and when omitting the longitudinal field component (thick solid line) for the spherical particle $x = 4$ and $m = 1.55$.

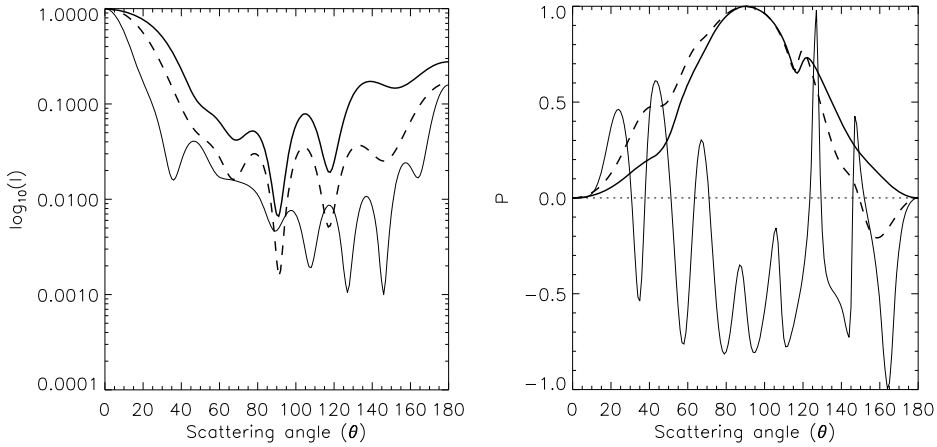


Fig. 11.5. Intensity $I_{\parallel}+I_{\perp}$ (left) and degree of linear polarization $P = (I_{\perp}-I_{\parallel})/(I_{\perp}+I_{\parallel})$ (right) in the case of the unmodified internal field (thin solid line), and when omitting the longitudinal field component and dividing the particle interior into either quadrants (thick dashed line) or sixteen cells (thick solid line) for the spherical particle with $x = 8$ and $m = 1.55$.

larization is positive for all the cases studied. This can be seen in Fig. 11.5 (right panel) for $x = 8$, and $m = 1.55$. The conclusion is that for spherical particles the longitudinal component seems to have a more pronounced influence on the general polarization, especially for the intermediate scattering angles, and that interference in the transverse component is responsible for both enhanced intensity and negative polarization near backscattering.

For G-sphere particles, Tyynelä et al. [34] have made use of a single sample G-sphere shape, ensemble-averaged size parameters equal to those of the previous study [33], and refractive indices $m = 1.55$ and $m = 1.33$. The relative standard deviation of radius has been $\sigma = 0.15$ and 0.245, and the power-law index of the covariance function has been $\nu = 4$. For $x = 2$, there are two large longitudinal-component maxima at the opposing sides of the particle. For the large ($x = 8$) G-sphere particle, there are two pairs of maxima for the longitudinal component. This is mostly due to the elongated shape of the particle, because the longitudinal component is focused on the forward part of the particle, which is similar to the internal-field structure for spherical particles. As size increases, this focusing concentrates more on the forward part. We estimate that, for very large particles, the effect of the longitudinal component on the linear polarization becomes less significant. Also, the irregular shape of the particle will further decrease the effect, because the longitudinal component will be weaker. When the longitudinal component is omitted, the degree of linear polarization changes to positive for all scattering angles except for $x > 6$. Negative polarization is still present for large G-sphere particles (Fig. 11.6 (right panel)). When the interior is divided into incoherently radiating quadrants, polarization becomes positive for all the cases studied (Fig. 11.7 (right panel)). Dividing the interior further into sixteen parts has little additional effect on the polarization. Only dividing the interior into incoherently radiating quadrants has a neutralizing effect for all scattering angles, especially

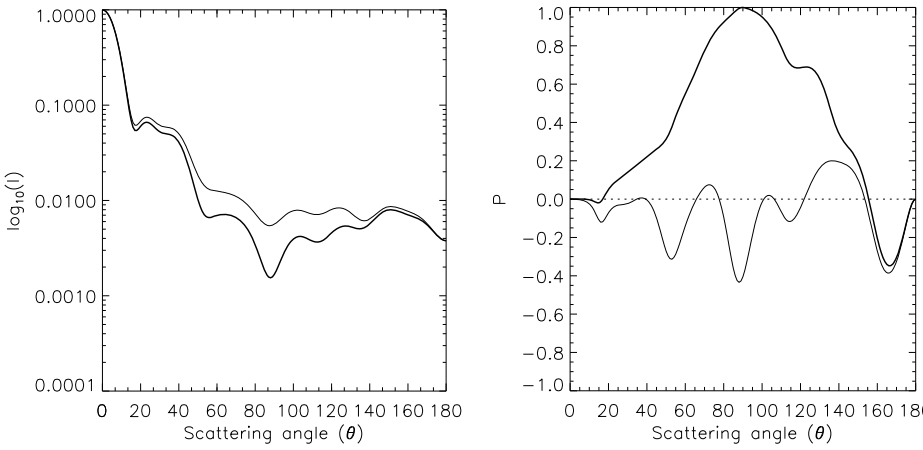


Fig. 11.6. Total intensity $I_{\parallel} + I_{\perp}$ (left) and degree of linear polarization $P = (I_{\perp} - I_{\parallel})/(I_{\perp} + I_{\parallel})$ (right) in the case of the unmodified internal field (thin solid line) and when the longitudinal component of the internal field is set to zero (thick solid line) for the G-sphere particle with $x = 8$, $m = 1.55$, and $\sigma = 0.15$.

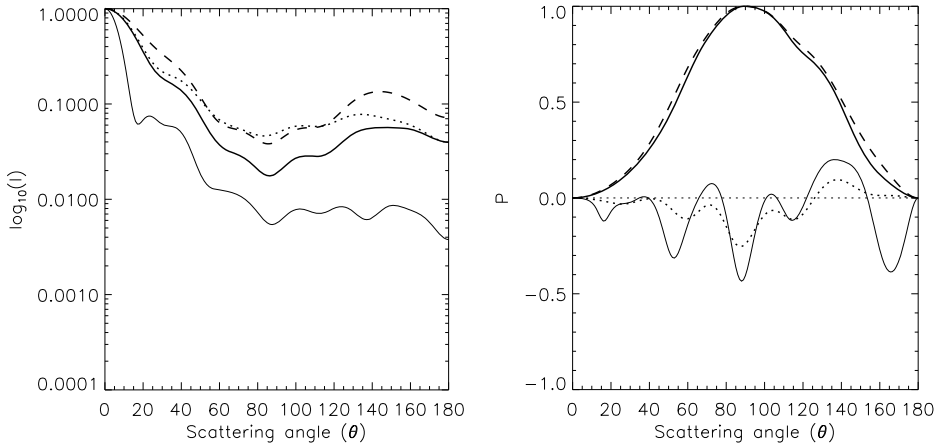


Fig. 11.7. Total intensity $I_{\parallel} + I_{\perp}$ (left) and degree of linear polarization $P = (I_{\perp} - I_{\parallel})/(I_{\perp} + I_{\parallel})$ (right) when the particle interior is divided into quadrants (thick dotted line), when the longitudinal component is omitted and the particle interior is divided into quadrants (thick solid line) or sixteen cells (thick dashed line), and when the internal field is unmodified (thin solid line). The parameters are the same as in Fig. 11.6.

near the backscattering direction, where the longitudinal component has little effect on polarization. Intensity enhancement seems to disappear when omitting only the transverse component. The conclusion is that for non-spherical particles, the interference effects appear to be weaker than for spherical particles due to their irregular shape.

11.2.4 Interference in single scattering

Zubko et al. [59] have carried out phase randomization in DDA calculations of light scattering from single dipoles to investigate interference effects inside irregularly shaped debris particles. They considered particles whose circumscribing-sphere size parameters $x_{\rm cs} = 6$ and $x_{\rm cs} = 10$ and refractive index $m = 1.5$. They have studied the interference effects of dipoles in five different cases: (1) the DDA solution (unmodified); (2) phase randomization between dipoles; (3) removal of the cross-interference between singly and multiply scattered waves; (4) removal of the interference of singly scattered waves; and, (5) the selection of mutually reciprocal waves. They have found that removing only the singly scattered waves (case 4) preserves negative polarization, but using phase randomization (case 2) or selecting only the reciprocal waves (case 5) makes negative polarization disappear. For $x = 6$, the influence of cross-interference (case 3) also makes negative polarization disappear. They conclude, first, that, in the case of single dipoles, the constructive interference between mutual reciprocal waves, the key constituent of the coherent-backscattering mechanism, cannot be responsible for negative polarization and intensity enhancement in the case studied. It is an open question at the moment whether or not groups of dipoles can be identified to give rise to coherent backscattering in the case of single particles.

Muinonen et al. [32] studied light scattering characteristics of G-sphere particles and suggested a single-scattering interference mechanism that can result in enhancement of intensity and negative polarization. The mechanism has two components, the longitudinal and transverse components that are responsible for the non-Rayleigh-like polarization features observed in single wavelength-scale particles.

In order to illustrate single-scattering interference of relevance to the present study [32, 33], we consider an electromagnetic plane wave (vector amplitude $\boldsymbol{E}_0$, wavelength λ, wave number $k = 2\pi/\lambda$, and wave vector $\boldsymbol{k}_0$) propagating along the z-axis and incident on a spherical scatterer located at the origin (Figs 11.8 and 11.9). Consider an observer in the xz-plane (scattering plane) with the scattering angle θ describing the angular deviation from the forward-scattering direction; thus, the phase angle is $\alpha = A = \pi - \theta$. In order to obtain the scattering characteristics for incident unpolarized light, the scattering problem needs to be solved for two perpendicular linear polarization states of the incident field, that is, for the y-polarized incident polarization vector perpendicular to the scattering plane and for the x-polarized incident polarization vector parallel to the scattering plane. The

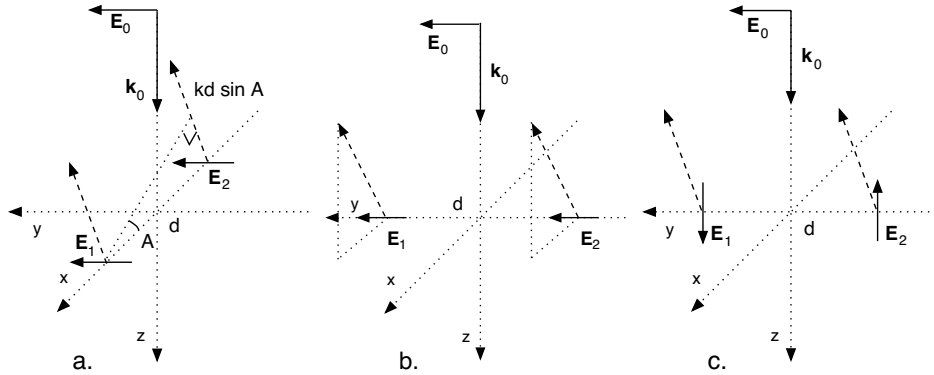


Fig. 11.8. Illustration of the induced internal-field components and their relation to the scattered field near the backward direction in the xz-plane for perpendicular incident polarization. See text for more details.

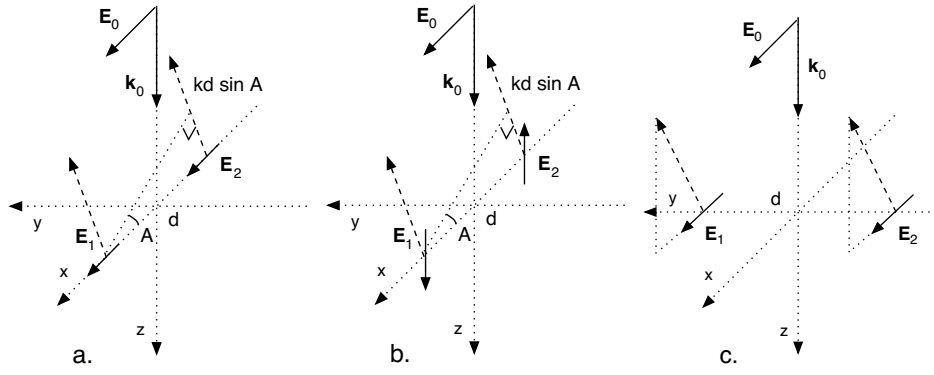


Fig. 11.9. As in Fig. 11.8 for parallel incident polarization.

final scattering characteristics follow as the average of the characteristics for the two polarization states of the incident wave. To further simplify the illustration, consider the internal fields induced in the spherical particle on the x and y-axes only. In reality, interference can take place at differing depths in the direction of the z-axis and the depths do not need to be equal among the pairs of locations shown in Figs 11.8 and 11.9. Assume for the time being that the six contributions depicted in Figs 11.8 and 11.9 do not interfere with each other.

First, consider an incident electric field polarized parallel to the y-axis corresponding to the perpendicular polarization state. On the x-axis (Fig. 11.8(a)), due to the symmetry of the particle, the incident wave gives rise only to internal-field components ($\boldsymbol{E}_1$ and $\boldsymbol{E}_2$ with subscripts referring to the two dipoles) perpendicular to the scattering plane, that is, a y-polarized internal field. Such an internal field gives rise only to a y-polarized or positively polarized scattered field in the xz-plane. Contributions to the scattered field from two mirror locations a distance d apart on the x-axis ($x > 0$ and $x < 0$) interfere constructively in the exact backward direction; whereas, the interference varies from constructive to destructive in other directions in the xz-plane. On the y-axis, due to the spherical symmetry, the incident y-polarized wave gives rise to both y-polarized (Fig. 11.8(b)) and z-polarized internal-field components (Fig. 11.8(c)). The z-components at the mirror locations of the y-axis ($y > 0$ and $y < 0$) have opposite signs but are otherwise equal. Because the phase difference equals π, the scattered field components arising from the z-components cancel each other. Note that, in the backward and forward directions, no contribution results from the z-polarized internal fields because there is no radiation along the line defined by the electric field vector of an electric dipole scatterer. The y-components on the two sides are equal and result in scattered wave components that interfere constructively for all scattering angles in the xz-plane.

Second, consider an incident electric field polarized parallel to the x-axis corresponding to the parallel polarization state. On the x-axis, due to the spherical symmetry, the incident x-polarized wave gives rise to both x-polarized (Fig. 11.9(a)) and z-polarized (Fig. 11.9(b)) internal-field components. The z-components at the two mirror locations of the x-axis ($x > 0$ and $x < 0$) have opposite signs but are otherwise equal, thus having a phase difference of π. The scattered-field components arising from the z-components typically interfere non-destructively with one another, giving rise to negative polarization across the entire scattering-angle regime except the exact backward and forward scattering directions where, again, no contribution follows from the z-components of the internal fields. On the y-axis (Fig. 11.9(c)), due to the spherical symmetry, the incident wave gives rise only to an x-polarized internal-field component. The x-components at the two mirror locations ($y > 0$ and $y < 0$) are equal and result in scattered wave components that interfere constructively for all scattering angles with zero contribution for the scattering angle of 90°.

The single-scattering mechanism is based on the hypothesis that, typically, for the incident y-polarized field, the y-polarized internal-field components on the x-axis are stronger than those on the y-axis and, similarly for the incident x-polarized field, the x-polarized internal-field components on the y-axis are stronger than those on the x-axis. As a net result close to backscattering, with the x-polarized internal-field components on the y-axis predominating over other contributions via

constructive interference, the degree of linear polarization for incident unpolarized light assumes negative values. This is enhanced by the first destructive interference geometry close to backscattering for the y-polarized scattered field arising from the x-axis. For both incident polarizations, the transverse internal-field components give rise to constructively interfering scattered-field components in the backward-scattering direction, resulting in a backscattering peak in the scattered intensity.

The single-scattering mechanism differs from the coherent-backscattering mechanism to be described in detail in section 11.3. In coherent backscattering, reciprocal waves travel through the same scatterers in a random medium, interfering constructively in the exact backscattering direction but not necessarily in other directions. In the single-scattering mechanism, the electric fields on the mirror locations describe the net result of all possible interactions among the electric dipoles constituting the scatterer. Last but not least, coherent backscattering is typically seen to occur in supermicron length scales; whereas, the single-scattering mechanism is relevant in submicrometer-to-micrometer length scales for visible light.

As stated above, the single-scattering mechanism has been verified for homogeneous spheres and G-sphere particles by Tyynelä et al. [33, 34]. Lumme et al. [76] and Mishchenko et al. [77] have reported increased backscattering and negative polarization near the backscattering direction for clusters of spherical particles. The angular widths of the features correlate with the size parameter of the clusters rather than with the extinction mean-free-path length of the media within the clusters. This dependence is qualitatively similar to the particle-size dependence of increased backscattering and negative polarization reported by Lumme and Rahola [37] and Muinonen et al. [42] for single-constituent Gaussian random particles. The latter features have been explained using the single-scattering interference mechanism (see sections 11.2.2, 11.2.3, and 11.2.4, as well as references therein). It remains as an open question which mechanism, CBM or the single-scattering mechanism, is responsible for the features reported for clusters of spherical particles.

11.2.5 Parameterizing single scattering

In what follows, we develop an analytical scattering parameterization based on pairs of electric dipoles [60]. We make use of Figs 11.8 and 11.9 in developing the amplitude scattering matrix elements and, in particular, fix the scattering plane to be the xz-plane. We follow the geometries in Figs 11.8 and 11.9 and assume that the dipoles are located either on the y-axis (scatterer 1) or on the x-axis (scatterer 2). Treated separately, these scatterers produce pure scattering matrices and well defined complex-valued amplitude-scattering matrices. We assess two perpendicular polarization states of the incident wave propagating in the positive direction of the z-axis. Our present modeling relies on the differing interference characteristics along the x-axis for the perpendicular and parallel polarizations.

Let us start by studying scatterer 1 constituting a system in the direction of the y-axis (Figs 11.8(b), 11.8(c), and 11.9(c)). Due to destructive interference, there is no contribution from the configuration in Fig. 11.8(c) in the scattering plane. Thus, the configuration in Fig. 11.8(b) is solely responsible for the contribution from the incident field perpendicular to the scattering plane. The configuration in Fig. 11.9(c) is solely responsible for the contribution from the incident field parallel

to the scattering plane. The amplitude scattering matrix is of a form identical to that of a Rayleigh particle:

$$S_1 = 2a_t \exp(\mathrm{i}\phi_t), \quad S_2 = 2a_t \exp(\mathrm{i}\phi_t)\cos\theta, \quad S_3 = S_4 = 0, \tag{11.5}$$

where a_t and ϕ_t are the amplitude and phase, respectively. Note, in particular, that there is no dependence on the distances between the dipoles. The Mueller scattering matrix coincides with that of the Rayleigh case and is independent of the absolute phase ϕ_t:

$$\begin{aligned} S_{11} = 2a_t^2(1+\cos^2\theta), \quad S_{12} = 2a_t^2(-1+\cos^2\theta), \quad S_{21} = S_{12}, \\ S_{22} = S_{11}, \quad S_{33} = 4a_t^2\cos\theta, \quad S_{44} = S_{33}, \end{aligned} \tag{11.6}$$

and the other elements being equal to zero.

Let us continue by studying scatterer 2, constituting a system in the direction of the x-axis (Figs 11.8(a), 11.9(a), and 11.9(b)). The configuration in Fig. 11.8(a) is responsible solely for the contribution from the incident field perpendicular to the scattering plane. However, the contributions from configurations in Figs 11.9(a) and 11.9(b) interfere with each other. For scatterer 2, we introduce the interdipole distances d_1, d_2, and d_3 corresponding to the configurations in Figs 11.8(a), 11.9(a), and 11.9b. Also, we assign different amplitudes and phase factors a_t, ϕ_t and a_l, ϕ_l for the transverse (Figs 11.8(a) and 11.9(a)) and longitudinal (Fig. 11.9(b)) electric dipoles, respectively. The amplitude scattering matrix takes the following form for scatterer 2:

$$\begin{aligned} S_1 &= 2a_t \exp(\mathrm{i}\phi_t)\cos\left(\frac{1}{2}kd_1\sin\theta\right), \\ S_2 &= 2a_t \exp(\mathrm{i}\phi_t)\cos\theta\cos\left(\frac{1}{2}kd_2\sin\theta\right) + \mathrm{i}2a_l\exp(\mathrm{i}\phi_l)\sin\theta\sin\left(\frac{1}{2}kd_3\sin\theta\right), \\ S_3 &= S_4 = 0. \end{aligned} \tag{11.7}$$

The Mueller scattering matrix follows from the amplitude matrix and is more complicated than that for scatterer 1:

$$\begin{aligned} S_{11} &= 2a_t^2\cos^2\left(\frac{1}{2}kd_1\sin\theta\right) + \\ &\quad 2a_t^2\cos^2\theta\cos^2\left(\frac{1}{2}kd_2\sin\theta\right) + 2a_l^2\sin^2\theta\sin^2\left(\frac{1}{2}kd_3\sin\theta\right) + \\ &\quad 4a_ta_l\cos\theta\sin\theta\cos\left(\frac{1}{2}kd_2\sin\theta\right)\sin\left(\frac{1}{2}kd_3\sin\theta\right)\sin(\phi_t-\phi_l), \\ S_{12} &= -2a_t^2\cos^2\left(\frac{1}{2}kd_1\sin\theta\right) + \\ &\quad 2a_t^2\cos^2\theta\cos^2\left(\frac{1}{2}kd_2\sin\theta\right) + 2a_l^2\sin^2\theta\sin^2\left(\frac{1}{2}kd_3\sin\theta\right) + \\ &\quad 4a_ta_l\cos\theta\sin\theta\cos\left(\frac{1}{2}kd_2\sin\theta\right)\sin\left(\frac{1}{2}kd_3\sin\theta\right)\sin(\phi_t-\phi_l), \end{aligned}$$

$$
\begin{aligned}
S_{21} &= S_{12}, \quad S_{22} = S_{11}, \\
S_{33} &= 4a_t^2 \cos\theta \cos\left(\frac{1}{2}kd_1 \sin\theta\right) \cos\left(\frac{1}{2}kd_2 \sin\theta\right) + \\
&\quad 4a_t a_l \sin\theta \cos\left(\frac{1}{2}kd_1 \sin\theta\right) \sin\left(\frac{1}{2}kd_3 \sin\theta\right) \sin(\phi_t - \phi_l), \\
S_{34} &= 4a_t a_l \sin\theta \cos\left(\frac{1}{2}kd_1 \sin\theta\right) \sin\left(\frac{1}{2}kd_3 \sin\theta\right) \cos(\phi_t - \phi_l), \\
S_{43} &= -S_{34}, \quad S_{44} = S_{33},
\end{aligned} \tag{11.8}
$$

and the other elements being equal to zero. Scatterer 2 introduces the possibility for a nonzero S_{34}-element, a clear deviation from the pure Rayleigh behavior of scatterer 1. Note that we have deliberately decided not to introduce different z-coordinates for the configurations in Figs 11.9(a) and 11.9(b). Such a difference would introduce additional phase differences that are left for a future study.

The amplitude scattering matrices for scatterers 1 and 2 can be developed further by assuming continuous lines of electric dipoles instead of two discrete dipoles. The amplitude scattering matrix for scatterer 1 remains unchanged; whereas, that for scatterer 2 is modified as follows where d_1 and d_2 now describe the half-lengths of the lines:

$$
\begin{aligned}
S_1 &= 2a_t \exp(\mathrm{i}\phi_t) \frac{\sin(\frac{1}{2}kd_1 \sin\theta)}{\frac{1}{2}kd_1 \sin\theta}, \\
S_2 &= 2a_t \exp(\mathrm{i}\phi_t) \cos\theta \frac{\sin(\frac{1}{2}kd_2 \sin\theta)}{\frac{1}{2}kd_2 \sin\theta} + \mathrm{i}2a_l \exp(\mathrm{i}\phi_l) \sin\theta \sin\left(\frac{1}{2}kd_3 \sin\theta\right), \\
S_3 &= S_4 = 0.
\end{aligned} \tag{11.9}
$$

Note that the longitudinal component remains discrete with inter-dipole distance d_3. The Mueller scattering matrix resembles that for scatterer 2 with discrete dipoles:

$$
\begin{aligned}
S_{11} &= 2a_t^2 \left[\frac{\sin(\frac{1}{2}kd_1 \sin\theta)}{\frac{1}{2}kd_1 \sin\theta}\right]^2 + \\
&\quad 2a_t^2 \cos^2\theta \left[\frac{\sin(\frac{1}{2}kd_2 \sin\theta)}{\frac{1}{2}kd_2 \sin\theta}\right]^2 + 2a_l^2 \sin^2\theta \sin^2\left(\frac{1}{2}kd_3 \sin\theta\right) + \\
&\quad 4a_t a_l \cos\theta \sin\theta \left[\frac{\sin(\frac{1}{2}kd_2 \sin\theta)}{\frac{1}{2}kd_2 \sin\theta}\right] \sin\left(\frac{1}{2}kd_3 \sin\theta\right) \sin(\phi_t - \phi_l), \\
S_{12} &= -2a_t^2 \left[\frac{\sin(\frac{1}{2}kd_1 \sin\theta)}{\frac{1}{2}kd_1 \sin\theta}\right]^2 + \\
&\quad 2a_t^2 \cos^2\theta \left[\frac{\sin(\frac{1}{2}kd_2 \sin\theta)}{\frac{1}{2}kd_2 \sin\theta}\right]^2 + 2a_l^2 \sin^2\theta \sin^2\left(\frac{1}{2}kd_3 \sin\theta\right) + \\
&\quad 4a_t a_l \cos\theta \sin\theta \left[\frac{\sin(\frac{1}{2}kd_2 \sin\theta)}{\frac{1}{2}kd_2 \sin\theta}\right] \sin\left(\frac{1}{2}kd_3 \sin\theta\right) \sin(\phi_t - \phi_l),
\end{aligned}
$$

$$
\begin{aligned}
S_{21} &= S_{12}, \;\; S_{22} = S_{11}, \\
S_{33} &= 4a_t^2 \cos\theta \left[\frac{\sin(\frac{1}{2}kd_1 \sin\theta)}{\frac{1}{2}kd_1 \sin\theta}\right] \left[\frac{\sin(\frac{1}{2}kd_2 \sin\theta)}{\frac{1}{2}kd_2 \sin\theta}\right] + \\
&\quad 4a_t a_l \sin\theta \left[\frac{\sin(\frac{1}{2}kd_1 \sin\theta)}{\frac{1}{2}kd_1 \sin\theta}\right] \sin\left(\frac{1}{2}kd_3 \sin\theta\right) \sin(\phi_t - \phi_l), \\
S_{34} &= 4a_t a_l \sin\theta \left[\frac{\sin(\frac{1}{2}kd_1 \sin\theta)}{\frac{1}{2}kd_1 \sin\theta}\right] \sin\left(\frac{1}{2}kd_3 \sin\theta\right) \cos(\phi_t - \phi_l), \\
S_{43} &= -S_{34}, \;\; S_{44} = S_{33},
\end{aligned}
\tag{11.10}
$$

and the other elements being equal to zero. As compared to the scattering matrix of scatterer 2 with two discrete electric dipoles, the one with the lines of dipoles produces smoother dependences, more closely mimicking the scattering characteristics of real particles.

We define the average single-scattering matrix as a weighted sum of those for the scatterers 1 and 2, multiplied by a function that allows us to obtain reasonable single-scattering phase functions:

$$
\boldsymbol{S} \;\propto\; f(\theta)\left[w_1 \boldsymbol{S}_1 + (1 - w_1)\boldsymbol{S}_2\right], \tag{11.11}
$$

where w_1 is the normalized weight of scatterer 1, $1 - w_1$ is the normalized weight of scatterer 2, and $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are the scattering matrices of scatterers 1 and 2, respectively. We require $w_1 \geq \frac{1}{2}$ in agreement with the natural assumption that there are equal numbers of scatterers 1 and 2 and that there are, in addition, pure Rayleigh-scatterers with scattering matrices equaling that of scatterer 1. Note that we do not provide any parameterization for the scattering cross section so Eq. (11.11) describes the shapes of the scattering matrix elements only.

For the function $f(\theta)$, we incorporate the double Henyey–Greenstein (2HG) phase function

$$
\begin{aligned}
f(\theta) &= w\frac{1 - g_1^2}{(1 + g_1^2 - 2g_1 \cos\theta)^{\frac{3}{2}}} + (1 - w)\frac{1 - g_2^2}{(1 + g_2^2 - 2g_2 \cos\theta)^{\frac{3}{2}}}, \\
g &= wg_1 + (1 - w)g_2,
\end{aligned}
\tag{11.12}
$$

where g_1 and g_2 describe the forward and backward asymmetries, w is the normalized weight of the first Henyey–Greenstein function, and g is the asymmetry parameter of the full 2HG phase function (see Fig. 11.13 for example 2HG phase functions). Note that, with the present multiplicative procedure, g is no longer the asymmetry parameter of the total scattering phase function. An alternative way to introduce the 2HG dependence is to make $f(\theta)$ contain in its denominator the S_{11} element of the combined scattering matrix of scatterers 1 and 2. The multiplication would then leave the pure 2HG function as the total scattering phase function. By choosing not to introduce the denominator, we are here conserving interference structures in the 11-element, too. For example scattering matrices, see Fig. 11.10. We show example coherent-backscattering computations using the parameterized scattering matrices in section 11.4.

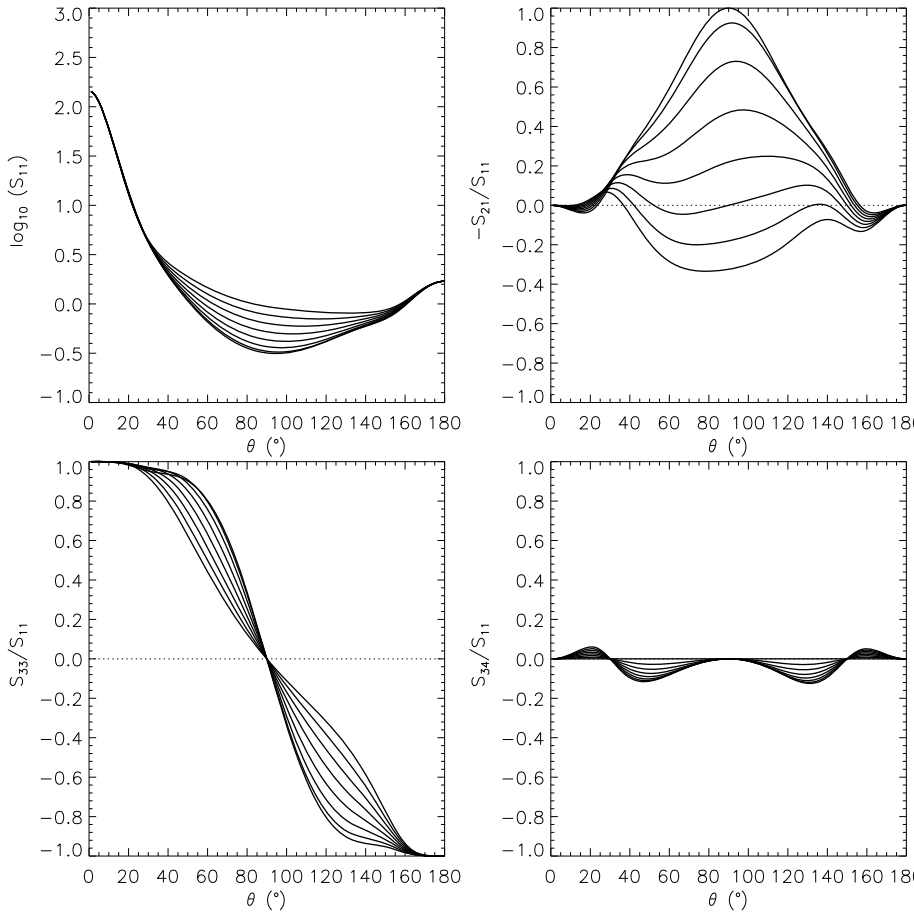


Fig. 11.10. Example analytical scattering-matrix elements (Eq. (11.11)) using the parameterization based on lines of dipoles (Eqs. (11.10) and (11.6)): $w = 0.8$, $g_1 = 0.8$, $g_2 = -0.2$, $w_1 = 0.5$, $a_t = 1$, $a_l = 0.0, 0.2, \ldots, 1.4$, $kd_1 = 4\pi$, $kd_2 = 3\pi$, $kd_3 = \pi$, and $\phi_t - \phi_l = \frac{7}{4}\pi$.

11.3 Coherent backscattering

Coherent backscattering by plane-parallel media of spherical scatterers has been treated earlier in [25] using numerical Monte Carlo methods, in which the coherent-backscattering contribution is computed alongside the radiative-transfer computation with the use of the reciprocity relation of electromagnetic scattering in the exact backscattering geometry. Typical for Monte Carlo methods, the modification of the original method for a spherical medium of scatterers is straightforward and established in the present chapter. The spherical media have, presently, infinite diametrical optical thicknesses, mimicking macroscopic objects.

11.3.1 Coherent-backscattering mechanism

CBM for the opposition effect is outlined in Fig. 11.11 for second-order scattering. An incident electromagnetic plane wave (solid and dashed lines; wavelength λ and wavenumber $k = 2\pi/\lambda$) interacts with two scatterers A and B before propagating to the observer on the left. We consider two components of the scattered wave, one which interacts with A first and then B and the reciprocal component that interacts with B first and then A. These two scattered wave components interfere constructively in directions defined by a cone constructed by rotating the light source direction vector S about the axis L joining the two scatterers. We illustrate a scattering direction on the cone precisely opposite to the light source direction; thus, the exact backward direction (phase angle $\alpha = 0°$) is on the constructive-interference cone for arbitrary locations of the two scatterers; whereas, in other directions, interference varies from constructive to destructive. Three-dimensional averaging over scatterer locations results in a net backscattering enhancement with decreasing angular width for increasing order of interactions, because the average distance between the end scatterers, like a random walk, tends to be larger for higher orders of interaction. The scattering processes can be caused by any disorder or irregularity in the medium.

CBM for the negative degree of linear polarization is explained for second-order scattering in Fig. 11.12. The incident radiation is unpolarized by definition, which requires the derivation and proper averaging of the Stokes vectors corresponding

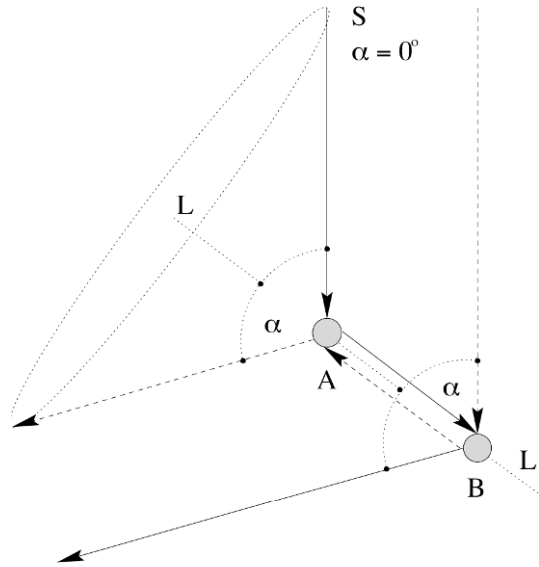


Fig. 11.11. Coherent backscattering mechanism (CBM) for the opposition effect. The multiply scattered electromagnetic wave components propagating in opposite directions between the scatterers (solid and dashed lines) interfere constructively in conical directions about the axis L, always including the exact backscattering direction (phase angle $\alpha = 0°$). In other directions, the interference is arbitrary depending on the wavelength, and the distance and orientation of the first and last scattering elements. See text.

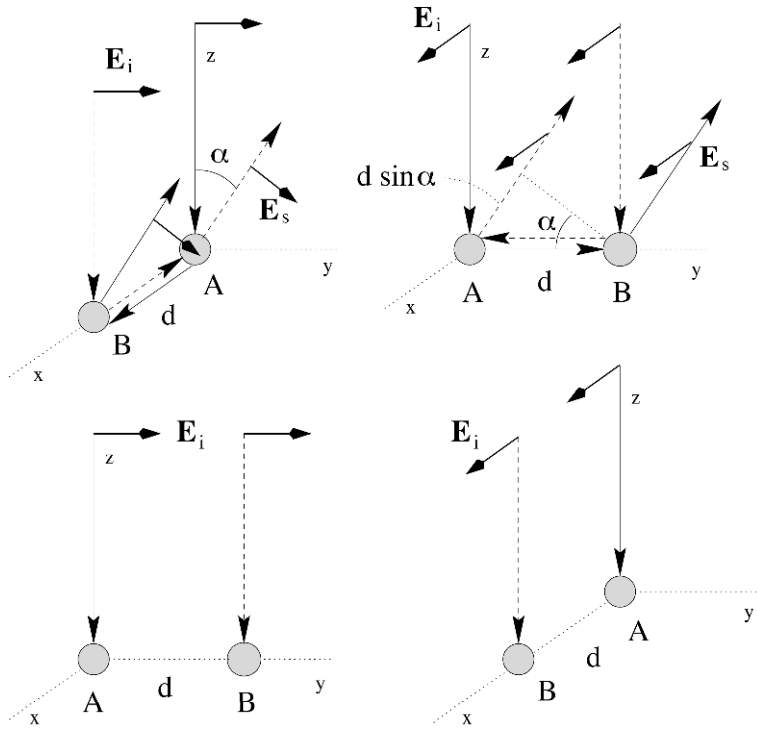


Fig. 11.12. Illustration of CBM for the negative linear polarization for unpolarized incident light. In the yz-plane in the scattering geometry leading to positive polarization, the interference depends on the phase difference $kd\sin\alpha$ (upper right panel); whereas, the interference is always constructive in the geometry causing negative polarization (upper left panel). Interaction with the polarization vector parallel to the line connecting the elements is typically suppressed (two lower panels) as compared to interaction with the perpendicular polarization vector. Averaging over scatterer locations results in an opposition effect and net negative polarization. See text.

to the scattered electromagnetic fields ($\boldsymbol{E}_\mathrm{s}$) for two linear polarization states of an incident plane wave ($\boldsymbol{E}_\mathrm{i}$). In Fig. 11.12, incident polarizations parallel and perpendicular to the scattering plane (here yz-plane) are treated in the two leftmost and two rightmost panels, respectively.

Consequently, in Fig. 11.12, an incident electromagnetic plane wave interacts with two scatterers A and B at a distance d from one another aligned either on the x-axis or the y-axis, while the observer is in the yz-plane. For the present geometries, the constructive interference cones of Fig. 11.11 reduce to the yz and xz-planes, depending on the alignment of the scatterers. Since first-order scattering is typically positively polarized (e.g., Rayleigh scattering and Fresnel reflection), the scatterers sufficiently far away from each other ($kd = 2\pi d/\lambda \gg 1$) interact predominantly with the electric field vector perpendicular to the plane defined by the source and the scatterers (two upper panels), while interaction with the electric field vector parallel to that plane is suppressed (two lower panels). The

observer in the yz-plane will detect negative polarization from the geometry in the upper left panel of Fig. 11.12, and positive polarization from the geometry in the upper right panel. However, the positive polarization suffers from the phase difference $kd\sin\alpha$; whereas the phase difference for the negative polarization is zero for all phase angles. Averaging over scatterer locations results in negative polarization near the backward direction. Scattering orders higher than the second experience similar preferential interaction geometries, and contribute to negative polarization. As above for the opposition effect, the contributions from increasing orders of scattering manifest themselves at decreasing phase angles.

Let us compare the interference in coherent backscattering with the interferences in single scattering in subsection 11.2.4. In coherent backscattering, reciprocal waves propagate in between discrete scatterers that are interacting with each other in a specific way. In single-scattering interference studies, we are dealing with internal fields of particles that are the result of all interactions within the particle. Whereas CBM is a unique mechanism for interference, a single universal mechanism in single-particle scattering may not necessarily exist.

The single-scattering studies so far have concentrated on identifying interference phenomena for single particles responsible for the single-particle intensity and polarization characteristics. On the other hand, if certain interference phenomena consistently repeat themselves for single particles, one may well consider incorporating the wording interference mechanism. Among the interference phenomena studied above, the longitudinal internal-field components have been seen to orchestrate the degree of linear polarization at intermediate scattering angles for all particles so far studied. We may thus claim the existence of a single-scattering interference mechanism for negative polarization at intermediate scattering angles.

Interference phenomena deriving from the transverse internal-field components arise from the superposition of contributions from varying locations inside the particles. It is too early to conclude if there is a leading contribution that manifests itself for all particles. At the moment, it is clear that interference phenomena in spherical particles differ from those in non-spherical particles. Moreover, it is impossible to pinpoint a single interference phenomenon responsible for the intensity and polarization patterns for all spherical particles near backscattering. In fact, for spherical particles, no independent transverse-field mechanism can exist [32, 33]: it has been obligatory to set the longitudinal component to zero before searching for transverse-field mechanisms.

For irregular single particles, there may well be a leading interference mechanism concerning transverse internal fields and resulting in enhanced backscattering and negative polarization near backscattering. There are studies under way to find such a mechanism. The ultimate mechanism, one may say, consists of Maxwell's equations of electromagnetism that dictate the scattering of light by small particles.

11.3.2 Theoretical framework for multiple scattering

For the physical quantities in multiple scattering, definitions in [61] are followed. The specific intensity of the incident radiation (in units of $\mathrm{J\,m^{-2}\,s^{-1}\,sr^{-1}}$) is

$$\boldsymbol{I}_0(\mu,\phi) \;=\; \pi\boldsymbol{F}_0\delta(\mu-\mu_0)\delta(\phi-\phi_0), \tag{11.13}$$

where $\pi \boldsymbol{F}_0$ is the incident flux density (in units of $\mathrm{J\,m^{-2}\,s^{-1}}$) and $\mu_0 = \cos\iota$. ι and ϕ_0 are the angle of incidence measured from the outward normal of the plane-parallel medium and the azimuthal angle of incidence, respectively.

The 4×4 reflection coefficient matrix $\boldsymbol{R}$ relates the incident flux density and the specific intensity of reflected radiation as

$$\begin{aligned} \boldsymbol{I}_{\mathrm{r}}(\mu,\phi) &= \mu_0 \boldsymbol{R}(\mu,\mu_0,\phi-\phi_0)\boldsymbol{F}_0, \\ \boldsymbol{R}(\mu,\mu_0,\phi-\phi_0) &= \boldsymbol{R}^{\mathrm{RT}}(\mu,\mu_0,\phi-\phi_0) + \boldsymbol{R}^{\mathrm{C}}(\mu,\mu_0,\phi-\phi_0), \end{aligned} \tag{11.14}$$

where $\mu = \cos\epsilon$. ϵ and ϕ are the angle of emergence measured from the outward normal of the plane-parallel medium and the azimuthal angle of emergence, respectively. The superscripts 'RT' and 'C' refer to the radiative-transfer and coherent-backscattering parts, respectively. The matrix element R_{11} corresponds to the so-called reflection coefficient in scalar radiative transfer. The enhancement factor is

$$\zeta(\mu,\mu_0,\phi-\phi_0) = \frac{R_{11}(\mu,\mu_0,\phi-\phi_0)}{R_{11}^{\mathrm{RT}}(\mu_0,\mu_0,0)}. \tag{11.15}$$

The degree of linear polarization is $P = -\tilde{R}_{21}/\tilde{R}_{11}$, where $\tilde{\boldsymbol{R}}$ is the reflection matrix expressed in the scattering plane defined by the source, object, and the observer.

The plane albedo $A_{\mathrm{p}}(\mu_0)$ is the fraction of the incident flux that is reflected by the plane-parallel medium in the case of incident unpolarized radiation,

$$A_{\mathrm{p}}(\mu_0) = \frac{1}{\pi}\int_0^1 \int_0^{2\pi} d\mu\, d\phi\, \mu R_{11}^{\mathrm{RT}}(\mu,\mu_0,\phi-\phi_0). \tag{11.16}$$

Note that the plane albedo here refers to the radiative transfer part only. The current treatment of energy conservation is sufficient for narrow coherent backscattering peaks.

Extinction is assumed to be exponential within the random medium. The optical depth τ_s between two locations a distance s apart is computed using the scalar extinction coefficient k_{e},

$$\tau_s = \int_0^s ds\, k_{\mathrm{e}} = k_{\mathrm{e}} s = \frac{s}{\ell}, \qquad \ell = \frac{1}{k_{\mathrm{e}}}, \tag{11.17}$$

where ℓ is the extinction mean free path. If n and v are the number and volume densities of spherical scatterers with radii a in a discrete random medium, we obtain

$$k_{\mathrm{e}} = n q_{\mathrm{e}} \pi a^2 = \frac{3 v q_{\mathrm{e}}}{4a}, \qquad n = \frac{3v}{4\pi a^3}. \tag{11.18}$$

In what follows, we assess coherent backscattering solely for spherical random media of scatterers. This is convenient, since we can characterize the resulting multiple scattering using the framework described in subsection 11.2.1 for single-particle scattering. The enhancement factor in Eq. (11.15) is readily generalized to the case of the spherical random medium. The so-called Bond albedo or spherical albedo A_{s} coincides with the single-particle albedo $\tilde{\omega}$. However, below, we do not incorporate coherent backscattering in the computation of A_{s} but give instead the value resulting from pure radiative transfer, $A_{\mathrm{s}} = A_{\mathrm{s}}^{\mathrm{RT}}$. The considerations of energy conservation are beyond the scope of the present work.

The disk-integrated brightness L (in units of $\mathrm{J\,s^{-1}\,sr^{-1}}$) equals the surface integral

$$\begin{aligned} L(\alpha) &= \int_{\mu,\mu_0>0} dA\ \mu I(\mu,\mu_0,\alpha) \\ &= \int_{\mu,\mu_0>0} dA\ \mu\mu_0 R_{11}(\mu,\mu_0,\alpha)F_0. \end{aligned} \tag{11.19}$$

For a spherical medium with diameter D and diametrical optical thickness approaching infinity so that the medium is locally plane parallel,

$$L(\alpha) = \frac{1}{4}D^2 \int_{\mu,\mu_0>0} d\Omega\ \mu\mu_0 R_{11}(\mu,\mu_0,\alpha)F_0. \tag{11.20}$$

The geometric albedo for the spherical random medium p is the ratio of the disk-integrated brightness of the medium (with diameter D) and the disk-integrated brightness of a normally illuminated Lambertian disk ($D^2\pi F_0/4$ with diameter D) at the exact backscattering direction $\alpha = 0°$:

$$p_{\mathrm{MS}} = \frac{L(0)}{L_{\mathrm{LD}}(0)} = 4\frac{L(0)}{D^2\pi F_0}. \tag{11.21}$$

Note that $0 < p < \infty$. For an irregularly shaped object (like an asteroid), the geometric albedo can depend on the orientation of the object. We distinguish between the geometric albedos including and excluding the coherent-backscattering contribution by denoting the geometric albedo from pure radiative transfer by p_{RT}. Thus, the enhancement factor gives the reflection coefficient divided by the pure radiative-transfer reflection coefficient at backscattering

$$\zeta(\alpha = 0) = \frac{p_{\mathrm{MS}}}{p_{\mathrm{RT}}}. \tag{11.22}$$

11.3.3 Scalar approximation

We present numerical computations for coherent backscattering of scalar waves by discrete scatterers with double Henyey–Greenstein single-scattering phase functions. The scatterers constitute spherical random media with radial optical thicknesses approaching infinity. We claim that the scalar approximation allows for an efficient initial assessment of multiple scattering. We recall that the physical explanations of the coherent-backscattering phenomenon in the intensity were based on the scalar approximation (e.g., [62–65]). Furthermore, Monte Carlo methods for scalar coherent backscattering were offered in [66] and subsequently applied to interpret the opposition effects of icy solar-system objects [67]. Recently, these methods have been utilized in the interpretation of the lunar imagery by AMIE (Advanced micro-Imager Experiment) onboard the SMART-1 mission to the Moon [68].

In the present Monte Carlo methods, coherent backscattering is computed alongside radiative transfer by incorporating the reciprocity relation of electromagnetic scattering in the backscattering geometry (e.g., [25]). The reciprocity

relation for scalar waves is more straightforward than that for vector waves. The numerical methods are optimized by the simultaneous computation of coherent backscattering for a large number of extinction mean free paths, resulting in an efficiency increased by a factor approaching the number of different mean free paths.

The present scalar studies rely on the double Henyey–Greenstein single-scattering phase function (2HG; see Eq. (11.12) and Fig. 11.13 for examples), allowing for a realistic modeling of the single-scattering phase functions. The multiple-scattering problem is thus formulated in terms of altogether five parameters: the single-scattering albedo $\tilde{\omega}$, the three 2HG parameters w, g_1, and g_2, as well as the extinction mean free path of the medium $k\ell$.

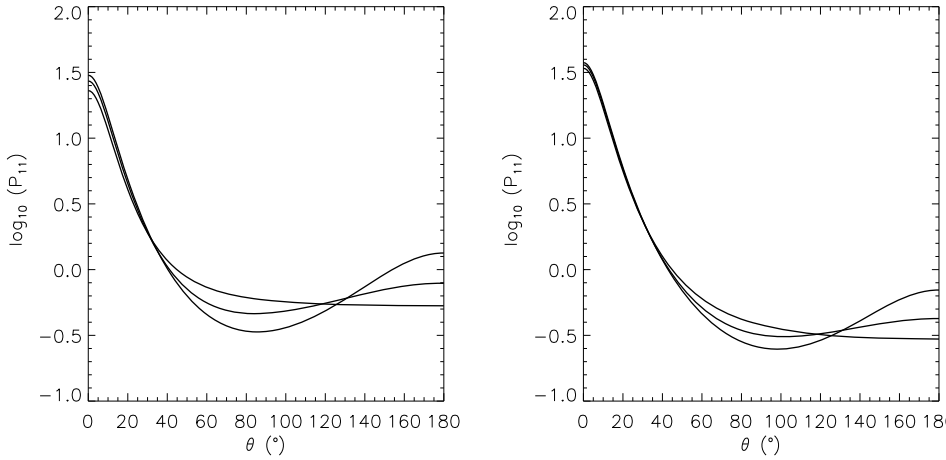


Fig. 11.13. Double Henyey–Greenstein phase functions for $g = 0.4$ (left) and $g = 0.6$ (right) with $g_1 = 0.8$ and $g_2 = -0.4$, -0.2, and 0.0.

We have carried out computations with 10 000 incident rays for a total of 20 832 spherical media of 2HG scatterers. The 16 single-scattering albedos are $\tilde{\omega} = 0.1$, 0.2, ..., 0.90, 0.95, 0.96, 0.97, 0.98, 0.99, 0.999, and 0.9999. The 62 dimensionless mean free paths are $k\ell = 2\pi\ell/\lambda = 30$, 40, 50, ..., 300, 400, 500, ..., 3000, 4000, 5000, ..., 10 000. The 21 2HG phase functions are $g = 0.1$, 0.2, ..., 0.7 when $g_1 = 0.8$ and $g_2 = -0.4$, -0.2, 0.0.

The enhancement factors ζ are shown in Fig. 11.14 in the case of $g = 0.6$, $g_1 = 0.8$, and $g_2 = -0.2$ for varying $\tilde{\omega}$ and $k\ell$. As is often observed for coherent backscattering, increasing $\tilde{\omega}$ and/or increasing $k\ell$ result in increasing $\zeta(0)$ and decreasing $\alpha_{\rm HWHM}$.

Fig. 11.15 shows the $\zeta(0)$ and $\alpha_{\rm HWHM}$ values against the geometric albedo of the spherical medium $p_{\rm MS}$. In Fig. 11.15(a), increasing $\tilde{\omega}$ is seen to result in increasing $\zeta(0)$ for three different 2HG phase functions with coinciding asymmetry parameter $g = 0.6$. In Fig. 11.15(b), decreasing $\alpha_{\rm HWHM}$ is seen to follow an increase in $\tilde{\omega}$.

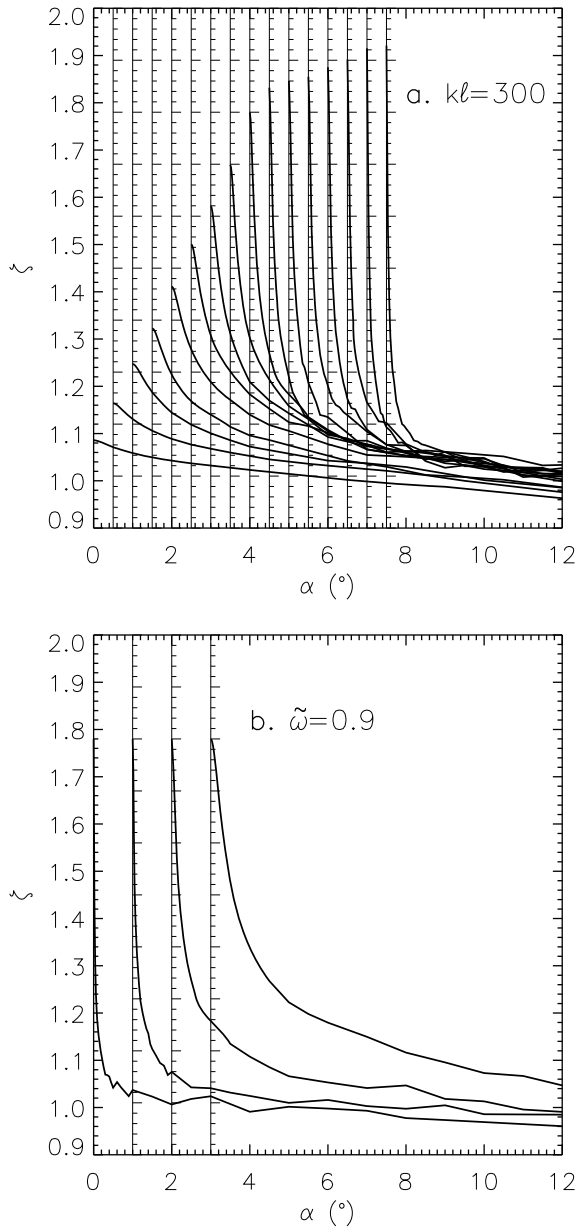


Fig. 11.14. Enhancement factors (ζ) for spherical media of scatterers (radial optical thickness $\tau_r \to \infty$) with double Henyey–Greenstein phase functions specified by $g = 0.6$, $g_1 = 0.8$, and $g_2 = -0.2$ for varying single-scattering albedos $\tilde{\omega}$ and extinction mean free paths $k\ell$: (a) $\tilde{\omega} = 0.1, 0.2, \ldots, 0.9, 0.95, 0.96, 0.97, 0.98, 0.99, 0.999, 0.9999$, $k\ell = 300$ (from left to right, shifted for better illustration); (b) $\tilde{\omega} = 0.9$, $k\ell = 100, 300, 1000, 3000$ (from left to right, shifted).

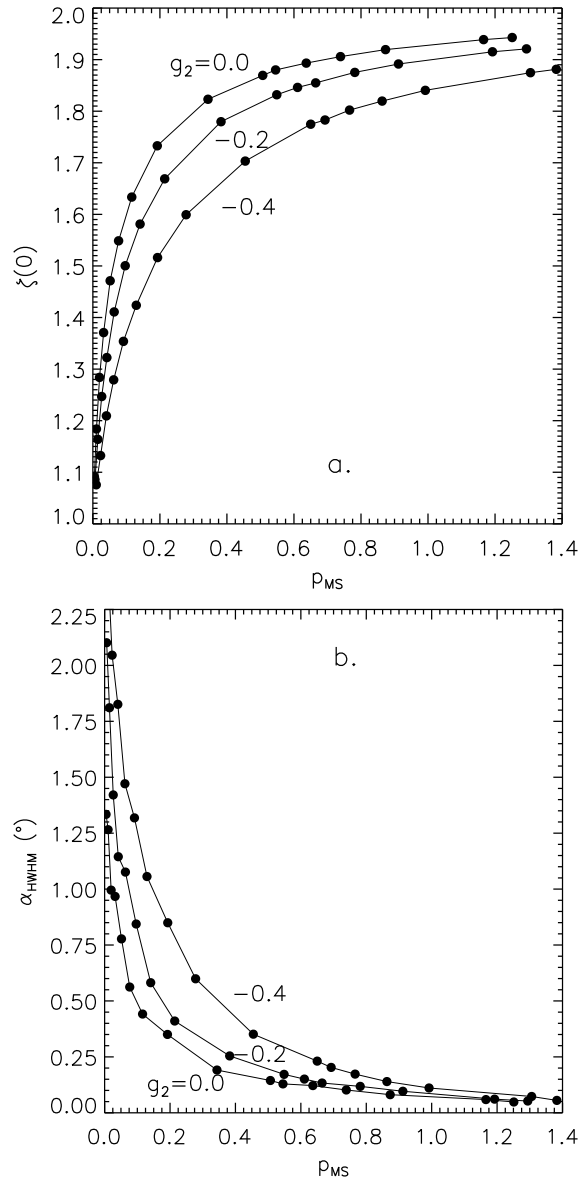


Fig. 11.15. Coherent backscattering by spherical media ($\tau_r \to \infty$) with $k\ell = 300$: (a) $\zeta(0)$ and (b) $\alpha_{\rm HWHM}$ vs. geometric albedo ($p_{\rm MS}$) for $g_1 = 0.8$, $g_2 = -0.4$, -0.2, 0.0, and $g = 0.6$, and varying $\tilde{\omega} = 0.1$, 0.2, ..., 0.7, 0.9, 0.95, 0.96, 0.97, 0.98, 0.99, 0.999, 0.9999 (from left to right).

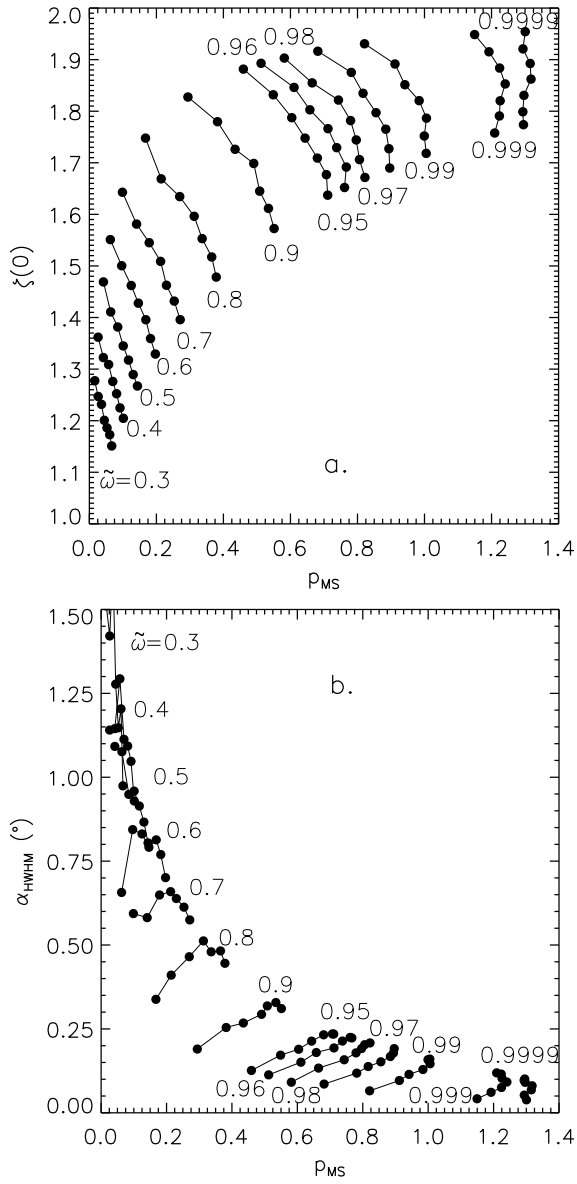


Fig. 11.16. Coherent backscattering by spherical media ($\tau_r \to \infty$) with $k\ell = 300$: (a) backscattering enhancement factors ($\zeta(0)$) and (b) angular widths ($\alpha_{\rm HWHM}$) vs. geometric albedo ($p_{\rm MS}$) for varying $\tilde{\omega}$, $g_1 = 0.8$, $g_2 = -0.2$, and $g = 0.1, 0.2, \ldots, 0.7$ ($\zeta(0)$ increases for increasing g).

The dependence of $\zeta(0)$ and $\alpha_{\rm HWHM}$ on g is illustrated in Fig. 11.16 for varying $\tilde{\omega}$. Based on Fig. 11.16(a), an increase in g results in a decrease in $p_{\rm MS}$ and an increase in $\zeta(0)$. This illustrates that small geometric albedos can result from random media with considerable contributions from multiple scattering. Figure 11.16(b) shows that increasing g decreases $\alpha_{\rm HWHM}$ (at least for higher $\tilde{\omega}$, cf. the accuracy of the present numerical computation). It is notable that $\zeta(0)$ and $\alpha_{\rm HWHM}$ depend on $\tilde{\omega}$ and the scattering phase function not only through the product $\tilde{\omega}g$: this can be ascertained by comparing $\zeta(0)$ and $\alpha_{\rm HWHM}$ for the cases $(\tilde{\omega}, g) = (0.6, 0.7)$ and $(\tilde{\omega}, g) = (0.7, 0.6)$ in Fig. 11.16.

We have compared the scalar and vector treatments for coherent backscattering by spherical media of infinite radial optical thickness in the case of Rayleigh scattering. For $\tilde{\omega} = 0.9$ and $k\ell = 300$, the enhancement factors over the pure radiative-transfer contribution at the phase angle $\alpha = 0^\circ$ are $\zeta(0) = 1.57$ and $\zeta(0) = 1.54$ for the vector and scalar cases, respectively. This difference is acceptable in many applications, warranting further studies of scalar coherent backscattering. We note that the deviation grows for increasing g as noticed in [20]. For a more extensive assessment of the scalar approximation, we refer the reader to the recent study by Mishchenko and Dlugach [69].

We have assessed the scalar approximation to coherent backscattering by optically thick spherical media of scatterers with the double Henyey–Greenstein phase functions. With the present multiple-scattering problem, we have mapped basic trends of coherent backscattering depending on the single-scattering albedo and phase function. The numerical results are available for further analyses and applications.

11.3.4 Vector approach

We present extensive numerical computations of coherent backscattering by spherical random media of Rayleigh scatterers. The radial optical thickness approaches infinity so that there are two parameters to be varied: the single-scattering albedo and the extinction mean free path. For each diffuse scattering process, the full scattering phase matrix is accounted for, as well as the full state of polarization of the incident Stokes vector. The scattering phase matrix can be specified using cubic splines for a user-defined matrix (cf., [25]).

Of interest in Monte Carlo multiple-scattering computations is the generation of scattering angles in each diffuse scattering process [25]. We denote the Stokes parameters of the incoming and outgoing rays by $\boldsymbol{I}_1 = (I_1, Q_1, U_1, V_1)^{\rm T}$ and $\boldsymbol{I}_2 = (I_2, Q_2, U_2, V_2)^{\rm T}$, respectively, and assume that the reference system of the incoming ray is arbitrary. The outgoing ray in the reference system of the local scattering plane is given by

$$\boldsymbol{I}_2 \;\propto\; \boldsymbol{P}(\theta, \phi) \cdot \boldsymbol{K}(\phi) \cdot \boldsymbol{I}_1, \tag{11.23}$$

where $\boldsymbol{K}$ is a 4×4 Mueller matrix for the rotation to the scattering plane, $\boldsymbol{P}$ is the 4×4 Mueller scattering matrix, and θ and ϕ are the temporary angles of scattering in the reference system of $\boldsymbol{I}_1$.

The Monte Carlo generation of the angles of scattering is based on the angular intensity distribution $I_2(\theta,\phi)$; i.e.,

$$\begin{aligned} I_2(\theta,\phi) \;\propto\; & I_1 P_{11}(\theta,\phi) + Q_1[P_{12}(\theta,\phi)\cos 2\phi + P_{13}(\theta,\phi)\sin 2\phi] + \\ & U_1[-P_{12}(\theta,\phi)\sin 2\phi + P_{13}(\theta,\phi)\cos 2\phi] + \\ & V_1 P_{14}(\theta,\phi). \end{aligned} \tag{11.24}$$

For scattering by spherical particles or ensemble-averaged Gaussian particles, the scattering matrix is independent of the azimuthal scattering angle ϕ and $P_{13} = P_{14} = 0$, so we obtain

$$I_2(\theta,\phi) \;\propto\; I_1 P_{11}(\theta) + Q_1 P_{12}(\theta)\cos 2\phi - U_1 P_{12}(\theta)\sin 2\phi. \tag{11.25}$$

The azimuthal angle ϕ shows up only in the rotation, and $I_2(\theta,\phi)$ is independent of the Stokes parameter V_1.

In order to obtain the probability density function $p(\theta,\phi)$, we normalize $I_2(\theta,\phi)$ and obtain

$$p(\theta,\phi) \;=\; \frac{1}{4\pi}P_{11}(\theta) + \frac{1}{4\pi}P_{12}(\theta)\left(\frac{Q_1}{I_1}\cos 2\phi - \frac{U_1}{I_1}\sin 2\phi\right). \tag{11.26}$$

The marginal probability density for θ is

$$p(\theta) \;=\; \frac{1}{2}P_{11}(\theta) \tag{11.27}$$

and thus independent of $\boldsymbol{I}_1$.

The practical generation of θ and ϕ can be carried out using, first, $p(\theta)$ to obtain the sample polar scattering angle $\tilde{\theta}$ and, second, the conditional probability

$$p(\phi \mid \tilde{\theta}) \;=\; \frac{p(\tilde{\theta},\phi)}{p(\tilde{\theta})} = \frac{1}{2\pi}\left[1 + \frac{P_{12}(\tilde{\theta})}{P_{11}(\tilde{\theta})}\left(\frac{Q_1}{I_1}\cos 2\phi - \frac{U_1}{I_1}\sin 2\phi\right)\right] \tag{11.28}$$

to obtain the sample azimuthal scattering angle $\tilde{\phi}$.

For the Rayleigh scattering matrix in Eq. (11.6), we obtain $\tilde{\theta}$-values analytically from

$$\cos\tilde{\theta} \;=\; \sqrt[3]{\sqrt{1+y^2}+y} - \sqrt[3]{\sqrt{1+y^2}-y}, \tag{11.29}$$

where y is a uniform random deviate $y \in \,]-2,2[$. For spherical scatterers beyond the Rayleigh regime, the generation of $\tilde{\theta}$ can best be carried out via interpolation in a precomputed table of $\cos\tilde{\theta}$-values as a function of $y \in \,]0,1[$.

For the subsequent generation of $\tilde{\phi}$, denote

$$\begin{aligned} e\cos\gamma \;&=\; -\frac{P_{12}(\tilde{\theta})}{P_{11}(\tilde{\theta})}\frac{Q_1}{I_1}, \\ e\sin\gamma \;&=\; -\frac{P_{12}(\tilde{\theta})}{P_{11}(\tilde{\theta})}\frac{U_1}{I_1}, \end{aligned} \tag{11.30}$$

where

$$e \;=\; \sqrt{\frac{P_{12}^2(\tilde{\theta})}{P_{11}^2(\tilde{\theta})}\,\frac{Q_1^2+U_1^2}{I_1^2}} \tag{11.31}$$

and the angle γ is unambiguously determined by Eq. (11.30). Then, the probability density function is

$$p(\phi \mid \tilde{\theta}) \;=\; \frac{1}{2\pi}\left[1 - e\cos(2\phi+\gamma)\right], \tag{11.32}$$

and the cumulative distribution function is

$$P(\phi \mid \tilde{\theta}) \;=\; \frac{1}{2\pi}\left[\phi - \frac{1}{2}e\sin(2\phi+\gamma) + \frac{1}{2}e\sin\gamma\right]. \tag{11.33}$$

For generating $\tilde{\phi}$, we thus obtain (with $y \in]0,1[$ a uniform random deviate)

$$(2\tilde{\phi}+\gamma) - e\sin(2\tilde{\phi}+\gamma) \;=\; 4\pi y + \gamma - e\sin\gamma, \tag{11.34}$$

that is, Kepler's equation $E - e\sin E = M$ with 'eccentric anomaly' $E = 2\tilde{\phi}+\gamma$, 'eccentricity' e, and 'mean anomaly' $M = 4\pi y + \gamma - e\sin\gamma$. Kepler's equation can be solved for $\tilde{\phi}$ efficiently using Newton's technique [71]: typically, only a few iterations are necessary to obtain $\tilde{\phi}$ with sufficient accuracy and the generation is nearly analytical.

The numerical methods have been optimized by introducing a simultaneous computation of coherent backscattering for a notable number of extinction mean free paths. This results again (cf. the scalar case) in an increased efficiency by a factor approaching that of the number of different mean free paths. There is only a small penalty in increasing the number of mean free paths in the computation.

We have coherent-backscattering computations carried out with 10 000 rays for a total of 630 spherical random media: the data base entails 21 single-scattering albedos $\tilde{\omega} = 0.05, 0.10, \ldots, 0.90, 0.95, 0.97, 0.99$, and 30 dimensionless mean free paths $k\ell = 2\pi\ell/\lambda = 10, 20, 30, \ldots, 100, 120, 140, \ldots, 200, 250, 300, 350, 400, 500, 600, \ldots, 1000, 2000, 3000, 4000, 5000, 10\,000$. The numerical database is available from the authors. Figs 11.17 and 11.18 show the main results from the numerical computations, and Table 11.1 gives the numerical values of the key parameters.

There remain several ways to optimize the coherent-backscattering computations for Rayleigh scatterers. First, coherent-backscattering contributions could be separately derived for each scattering order higher than the second (inclusive; up to the order of 10,000), allowing a rapid application for user-specified single-scattering albedos. Second, in principle, the extinction mean free paths could all be treated simultaneously by separating the mean free paths $k\ell$ and angular factors in the relevant electromagnetic phases during the computation. This would require the development of sinusoidal functions with product-type arguments into presentations where the mean-free-path parts were separated from the geometrical parts. Third, symmetries could be utilized in the Monte Carlo computation allowing for improved convergence characteristics for polarization, in particular.

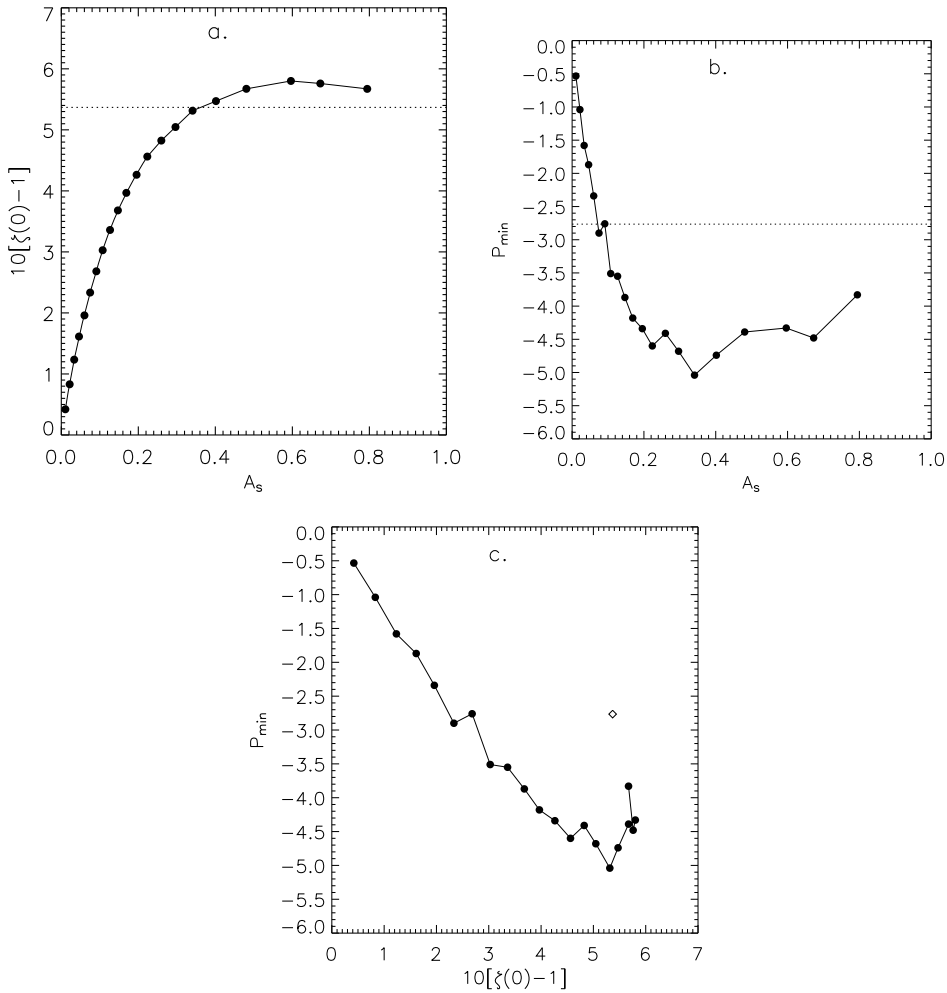


Fig. 11.17. Enhancement factors in the backscattering direction ($\zeta(0)$) and minimum polarizations ($P_{\min}$) for spherical random media of infinite diametrical optical thickness with $k\ell = 300$ consisting of Rayleigh scatterers with varying single-scattering albedos $\tilde{\omega}$. Note the differing turning points for $\zeta(0)$ and $P_{\min}$ and the literature values [70] for a conservative plane-parallel medium at normal incidence ($\tilde{\omega} = 1.0$: $\zeta(0) = 1.5368$, $P_{\min} = -2.765\%$): (a) $10[\zeta(0) - 1]$ vs. $A_{\rm s}$, (b) $P_{\min}$ vs. $A_{\rm s}$, and (c) $P_{\min}$ vs. $10[\zeta(0) - 1]$.

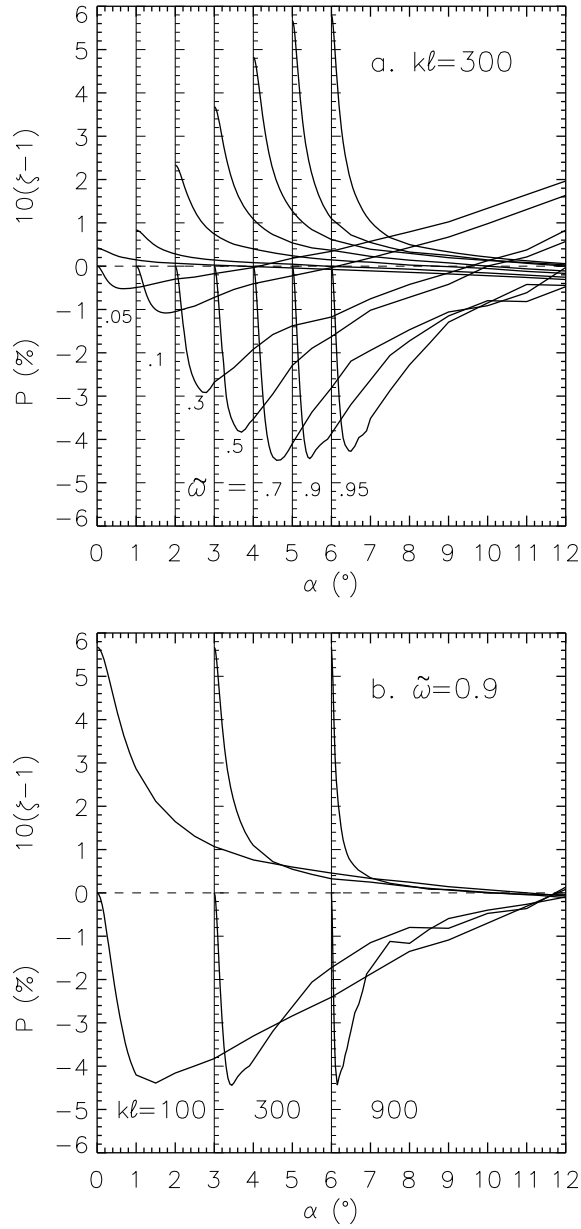


Fig. 11.18. Enhancement factors and polarization surges for spherical random media of Rayleigh scatterers with varying single-scattering albedos $\tilde{\omega}$ and extinction mean free paths $k\ell$ (normal incidence): (a) $\tilde{\omega} = 0.05$, 0.1, 0.3, 0.5, 0.7, 0.9, 0.95 and $k\ell = 300$; (b) $\tilde{\omega} = 0.9$ and $k\ell = 100$, 300, 900. Angular scales shifted for better illustration.

Table 11.1. Spherical albedos A_s (%), geometrical albedos with and without the coherent-backscattering peak p_{MS} and p_{RT}, respectively, enhancement factors $\zeta(0)$, and minimum polarizations P_{min} (%) and phase angles of minimum polarization α_{min} (°) for spherical media of infinite diametrical optical thickness consisting of Rayleigh scatterers with single-scattering albedo $\tilde{\omega}$. For the minimum polarization, the extinction mean free path is $k\ell = 300$.

$\tilde{\omega}$	A_s	p_{MS}	p_{RT}	$\zeta(0)$	P_{min}	α_{min}
0.05	1.05	0.00681	0.00654	1.042	-0.534	0.7
0.10	2.18	0.0148	0.0137	1.083	-1.04	0.8
0.15	3.35	0.0240	0.0214	1.123	-1.58	0.9
0.20	4.60	0.0345	0.0297	1.161	-1.87	0.5
0.25	6.02	0.0467	0.0391	1.196	-2.34	0.8
0.30	7.49	0.0607	0.0492	1.233	-2.90	0.7
0.35	9.10	0.0769	0.0607	1.268	-2.76	0.9
0.40	10.7	0.0946	0.0726	1.303	-3.51	0.8
0.45	12.7	0.115	0.0861	1.336	-3.55	0.5
0.50	14.7	0.139	0.101	1.368	-3.87	0.7
0.55	16.9	0.164	0.118	1.397	-4.18	0.7
0.60	19.6	0.196	0.137	1.426	-4.34	0.7
0.65	22.3	0.230	0.158	1.456	-4.60	0.6
0.70	26.0	0.274	0.185	1.482	-4.41	0.7
0.75	29.7	0.321	0.213	1.505	-4.68	0.5
0.80	34.1	0.379	0.247	1.531	-5.04	0.5
0.85	40.2	0.451	0.291	1.547	-4.74	0.5
0.90	48.1	0.551	0.352	1.567	-4.39	0.45
0.95	59.6	0.692	0.438	1.580	-4.33	0.5
0.97	67.3	0.780	0.495	1.576	-4.48	0.45
0.99	79.4	0.916	0.585	1.567	-3.83	0.5

11.4 Physical modeling

11.4.1 Polarization fits

The scattering parameterization for single scattering described above in subsection 11.2.5 allows for the development of a polarization model that can be directly utilized in the classification of the polarimetric observations of atmosphereless solar-system objects. Consider the amplitude scattering matrices for scatterers 1 and 2 presented in subsection 11.2.5. We divide the matrix elements S_1 and S_2 for scatterers 1 and 2 with the amplitude term $2a_t \exp(\mathrm{i}\phi_t)$. For scatterer 1, we obtain

$$S_1 = 1, \quad S_2 = \cos\theta, \quad S_3 = S_4 = 0; \tag{11.35}$$

whereas, for scatterer 2, we have

$$S_1 = \frac{\sin(\frac{1}{2}kd_1\sin\theta)}{\frac{1}{2}kd_1\sin\theta}, \qquad S_2 = \cos\theta\frac{\sin(\frac{1}{2}kd_2\sin\theta)}{\frac{1}{2}kd_2\sin\theta} + Z\sin\theta\sin\left(\frac{1}{2}kd_3\sin\theta\right),$$
$$S_3 = S_4 = 0, \tag{11.36}$$

where the complex-valued parameter

$$Z \;=\; \mathrm{i}\frac{a_l}{a_t}\exp\left[\mathrm{i}(\phi_l-\phi_t)\right]. \tag{11.37}$$

We introduce averaging over the lengths of the lines of electric dipoles, mimicking integration over a size distribution of scatterers in the regolith. For scatterer 1, the length-averaged scattering matrix equals the Rayleigh scattering matrix, as there is no dependence on line length in $\boldsymbol{S}_1$:

$$\boldsymbol{F}_1 \;=\; \int_0^\infty\int_0^\infty du_1\,du_2\,p(u_1,u_2)\boldsymbol{S}_1 = \boldsymbol{S}_1, \tag{11.38}$$

where u_1 and u_2 stand for the integration variables corresponding to kd_1 and kd_2, respectively, and $p(u_1,u_2)$ is the joint probability density function (p.d.f.) for u_1 and u_2. For scatterer 2, size-averaging starts from

$$\boldsymbol{F}_2 \;=\; \int_0^\infty\int_0^\infty du_1\,du_2\,p(u_1,u_2)\boldsymbol{S}_2. \tag{11.39}$$

In what follows, we consider lines of dipoles; whereas, it would be completely feasible to consider discrete dipoles in an analogous manner. Assuming that u_1 and u_2 are independent, $p(u_1,u_2)=p(u_1)p(u_2)$ and the integrals reduce to the computation of integrals of the type

$$I_n \;=\; \int_0^\infty du\;p(u)\left[\frac{\sin(\frac{1}{2}u\sin\theta}{\frac{1}{2}u\sin\theta}\right]^n \tag{11.40}$$

that, in general, require numerical integration that can be carried out using, e.g., the Gauss–Legendre integration scheme [72].

The inverse-gamma distribution is an attractive choice for averaging, due to its power-law asymptotics for large arguments:

$$p(u) \;=\; \frac{y^s}{\Gamma(s)}\frac{1}{u^{s+1}}\exp\left(-\frac{y}{u}\right), \tag{11.41}$$

where Γ is the gamma function. Here y is called the scale parameter and s the shape parameter of the distribution. Asymptotically, $s+1$ corresponds to the power law index of the size distribution. For a given wavelength of incident light, the scale parameter y provides a length measure in the regolith in the size-parameter domain; that is, if k is the wave number and $kd=y$, d is the length corresponding to y. There is no *a priori* reason why the length distributions would be the same for kd_1 and kd_2, resulting in altogether four parameters (y_1,s_1) and (y_2,s_2) describing the characteristics of the line sources.

For scatterer 2 in Eq. (11.36), we have still the parameters Z and kd_3 to assess. In introducing the length averaging above for kd_1 and kd_2, we have assumed distributions independent from one another. Due to the more complicated nature of Z, rather than introducing averaging, we treat its real and imaginary parts $\mathrm{Re}Z$ and $\mathrm{Im}Z$ as free parameters of the physical model. For kd_3, we make use of the value $kd_3=\pi$, guided by our studies on the internal fields of spherical particles in subsection 11.2.3.

The physical model for the degree of linear polarization yields

$$P(\theta) \;=\; \frac{-w_1 F_{21}^{(1)} - (1-w_1) F_{21}^{(2)}}{w_1 F_{11}^{(1)} + (1-w_1) F_{11}^{(2)}}, \tag{11.42}$$

where w_1 and $1 - w_1$ are the normalized weights of scatterers 1 and 2. As earlier, we require $w_1 \geq \frac{1}{2}$. The physical model depends on seven free parameters (fixing $kd_3 = \pi$): w_1, (y_1, s_1), (y_2, s_2), Re Z, and Im Z.

In Fig. 11.19, we show the polarimetric observations of asteroid (1) Ceres [73, 74] explained in terms of the physical model using Markov-Chain Monte Carlo sampling for the model parameters. Assuming $s_1 = s_2$, the best-fit model parameters with 3-σ (or 99.7%) errors are as follows: $w_1 = 0.808$ (+0.007, −0.009), $s_1 = s_2 = 0.73$ (+0.05, −0.02), $y_1 = 3.3$ (+0.2, −0.2), $y_2 = 0.205$ (+0.002, −0.016), Re$Z = 0.0037$ (+0.0383, −0.9993), Im$Z = 2.06$ (+0.10, −0.01). The rms-value of the fit is 0.068%. The 3-σ error envelope has been computed using 10 000 sample solutions for the parameters and finding a maximum and minimum among different model predictions at each phase angle. The physical model provides an acceptable fit to the polarimetric observations. The model parameters are all concentrated in the proximity of their best-fit values. Furthermore, successful fitting with $s_1 = s_2$ indicates consistency of the distributions for kd_1 and kd_2. The shape of the negative polarization branch results from the considerably deviating scale parameters y_1 and y_2. The shape differs from that in the single-scattering modeling, where realistic scattering matrices for wavelength-scale scatterers follow from kd_1 and kd_2 only slightly differing from one another.

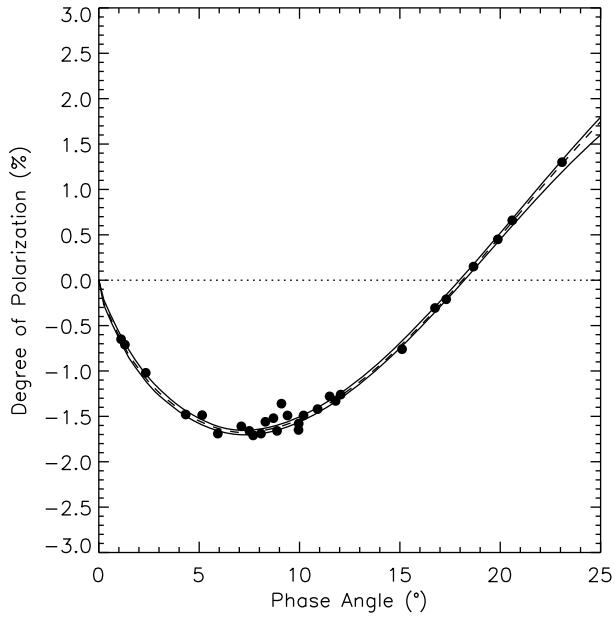


Fig. 11.19. Physical model for the polarimetry of asteroid (1) Ceres.

11.4.2 Coherent-backscattering simulations

We carry out example coherent-backscattering simulations for infinitely thick spherical random media of scatterers composed of lines of dipoles. These simulations are required for the simultaneous interpretation of photometric and polarimetric observations for atmosphereless solar-system objects. The vector treatment removes the shortcomings of the scalar approximation discussed in subsection 11.3.3. Before entering the simulations in detail, we refer the reader to the coherent-backscattering calculations by Videen et al. [75] for comparison.

In the simulations, we utilize parameters indicated in Fig. 11.10, choosing $a_l = 0.6$, for three single-scattering albedos $\tilde{\omega} = 0.3$, 0.6, and 0.9 and three extinction mean free paths $k\ell = 100$, 300, and 1000. For the cases $\tilde{\omega} = 0.3$, 0.6, and 0.9, we sample 200 000, 200 000, and 100 000 incident rays.

Figure 11.20 shows the photometric and polarimetric phase functions for the spherical media for a number of mean free paths ($k\ell = 100$, 300, and 1000). Here ζ is the enhancement factor compared to the pure radiative-transfer contribution at the exact backscattering direction (e.g., [25]), and $P = (I_\perp - I_\parallel)/(I_\perp + I_\parallel)$ is the degree of linear polarization for unpolarized incident light (the subscripts referring to the intensity components perpendicular and parallel to the scattering plane). The single-scattering characteristics are seen to be neutralized and coherent backscattering gives rise to sharp photometric and polarimetric surges close to the backscattering direction. A double-minimum feature appears for $\tilde{\omega} = 0.6$ in Fig. 11.20: it is straightforward to identify the negative coherent-backscattering and single-scattering lobes in the polarization curve. Note that, for $\tilde{\omega} = 0.9$, the single-scattering lobe neutralizes due to the dominating role of multiple scattering.

With the help of parameterized single scattering in coherent-backscattering modeling, it is possible to qualitatively explain the narrow opposition effects and negative polarization branches for E-class asteroids by Rosenbush et al. [4, 5]. Furthermore, the peculiar polarization curve observed for asteroid (234) Barbara by Cellino *et al.* [9] resembles, to some extent, the result in Fig. 11.20. Note that it is straightforward to incorporate distributions of empirical scatterers in the coherent-backscattering computation, allowing for a realistic coherent-backscattering modeling of photometric and polarimetric phase effects.

11.5 Conclusion

We have assessed light-scattering problems relevant for the interpretation of the photometric and polarimetric observations of atmosphereless solar-system objects. In single scattering of light by small particles, we have studied intensity and polarization phenomena at both intermediate scattering angles and close to backscattering. We have described interference mechanisms at least partly responsible for the phenomena, focusing on the interrelation of the internal and scattered fields of the particles. We have developed an analytical scattering-matrix parameterization for single scattering meant to be utilized in solving inverse scattering problems. In multiple scattering of light by complex random media of scatterers, we have studied coherent backscattering using the scalar approximation omitting polarization as well as following the rigorous vector approach accounting for polarization.

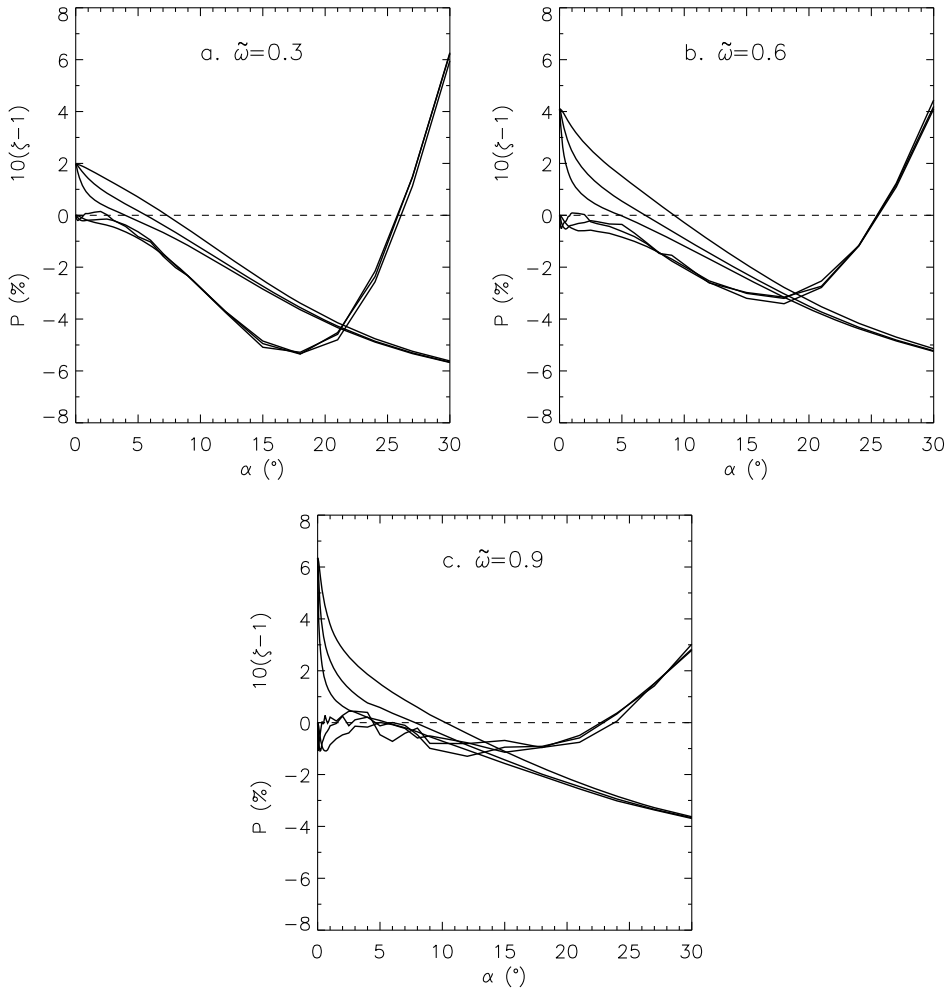


Fig. 11.20. Example coherent-backscattering results for spherical random media of infinite optical thickness with mean free paths $k\ell = 100$, 300, and 1000: (a) $\tilde{\omega} = 0.3$; (b) $\tilde{\omega} = 0.6$; (c) $\tilde{\omega} = 0.9$.

We have concentrated on spherical media of scatterers, allowed by the flexible Monte Carlo algorithm for numerical computations of coherent backscattering. In the coherent-backscattering studies, we have utilized the scattering-matrix parameterization earlier described for single scattering. Finally, we have described practical physical modeling that can be applied in analyses of the polarimetric and photometric observations of solar-system objects.

Acknowledgments

KM wishes to thank J. Peltoniemi, H. Parviainen, and J. Näränen for numerous discussions on multiple scattering by planetary regoliths. The authors appreciate the critical remarks by the two anonymous referees. Research supported, in part, by the Academy of Finland (project No. 127461).

References

1. T. Gehrels. Photometric studies of asteroids. V. The light-curve and phase function of 20 Massalia: *Astrophys. J.* **123**, 331–338, 1956.
2. B. Lyot. Recherches sur la polarisation de la lumière des planètes et de quelques substances terrestres: *Ann. Obs. Paris* **8**, 1–161, 1929.
3. A. W. Harris, J. W. Young, L. Contreiras, T. Dockweiler, L. Belkora, H. Salo, W. D. Harris, E. Bowell, M. Poutanen, R. P. Binzel, D. J. Tholen, S. Wang. Phase relations of high albedo asteroids: The unusual opposition brightening of 44 Nysa and 64 Angelina: *Icarus* **81**, 365–374, 1989.
4. V. K. Rosenbush, N. N. Kiselev, V. G. Shevchenko, K. Jockers, N. M. Shakhovskoy, Y. S. Efimov. Polarization and brightness opposition effects for the E-type asteroid 64 Angelina: *Icarus* **178**, 222–234, 2005.
5. V. K. Rosenbush, V. G. Shevchenko, N. N. Kiselev, A. V. Sergeev, N. M. Shakhovskoy, F. P. Velichko, S. V. Kolesnikov, N. V. Karpov. Polarization and brightness opposition effects for the E-type Asteroid 44 Nysa: *Icarus* **201**, 655–665, 2009.
6. H. Boehnhardt, S. Bagnulo, K. Muinonen, M. A. Barucci, L. Kolokolova, E. Dotto, G. P. Tozzi. Surface characterization of 28978 Ixion (2001 KX_{76}): *Astron. Astrophys.* **415**, L21–L25, 2004.
7. S. Bagnulo, H. Boehnhardt, K. Muinonen, L. Kolokolova, I. Belskaya, M. A. Barucci. Exploring the surface structure of transneptunian objects and Centaurs with polarimetric FORS1/VLT observations. *Astron. Astrophys.* **450**, 1239–1248, 2006.
8. P. Rousselot, A.-C. Levasseur-Regourd, K. Muinonen, J.-M. Petit. Polarimetric and photometric phase effects observed on transneptunian object (29981) 1999 TD_{10}: *Earth, Moon, and Planets* **97**(3-4), 353–364, 2005.
9. A. Cellino, I. N. Belskaya, P. Bendjoya, M. Di Martino, R. Gil-Hutton, K. Muinonen, and E. F. Tedesco. The strange polarimetric behaviour of asteroid (234) Barbara: *Icarus* **180**, 565–567, 2006.
10. I. N. Belskaya, A.-C. Levasseur-Regourd, Y. G. Shkuratov, K. Muinonen. Surface properties of Kuiper-Belt objects and Centaurs from photometry and polarimetry. In *The Solar System Beyond Neptune* (M. A. Barucci, H. Boehnhardt, D. P. Cruikshank, and A. Morbidelli, Eds., University of Arizona Press, Tucson, Arizona, U.S.A.), pp. 115–127, 2008.
11. K. Muinonen, J. Piironen, Yu. G. Shkuratov, A. Ovcharenko, B. Clark. Asteroid photometric and polarimetric phase effects. In *Asteroids III* (W. Bottke, R. P. Binzel, A. Cellino, and P. Paolicchi, Eds, Tucson: University of Arizona Press), pp. 123–138, 2002.
12. V. Rosenbush, N. Kiselev, V. Avramchuk, M. Mishchenko. Photometric and polarimetric phenomena exhibited by solar system bodies. In *Optics of Cosmic Dust* (G. Videen and M. Kocifaj, Eds, Dordrecht: Kluwer Academic Publishers), pp. 191–224, 2002.
13. B. Hapke. Bidirectional reflectance spectroscopy: 4. The extinction coefficient and the opposition effect: *Icarus* **67**, 264–280, 1986.

14. K. Lumme, E. Bowell. Radiative transfer in the surfaces of atmosphereless bodies. I. Theory: *Astron. J.* **86**, 1694–1704, 1981.
15. Y. G. Shkuratov, M. A. Kreslavsky, A. A. Ovcharenko, D. G. Stankevich E. S. Zubko, C. Pieters, G. Arnold. Opposition effect from Clementine data and mechanisms of backscatter: *Icarus* **141**, 132–155, 1999.
16. Y. G. Shkuratov. On the origin of the opposition effect and negative polarization for cosmic bodies with solid surface: *Astron. Circular No. 1400*, Sternberg State Astron. Inst., Moscow, 3–6, 1985.
17. K. Muinonen. Electromagnetic scattering by two interacting dipoles: *Proc. 1989 URSI Electromagnetic Theory Symposium* (Stockholm) 428–430, 1989.
18. K. Muinonen. Light scattering by inhomogeneous media: backward enhancement and reversal of polarization, PhD-thesis, University of Helsinki, 1990.
19. B. Hapke. Coherent backscatter and the radar characteristics of outer planet satellites: *Icarus* **88**, 407–417, 1990.
20. M. I. Mishchenko, J. M. Dlugach. Coherent backscatter and the opposition effect for E-type asteroids: *Planet. Space Sci.* **41**, 173–181, 1993.
21. M. Wolff. Polarization of light reflected from rough planetary surface: *Appl. Opt.* **14**, 1395–1405, 1975.
22. M. I. Mishchenko, L. D. Travis, A. A. Lacis: *Multiple Scattering of Light by Particles: Radiative Transfer and Coherent Backscattering*, Cambridge University Press, 2006.
23. K. Muinonen. Coherent backscattering by solar system dust particles. In *IAU Symposium No. 160, Asteroids, Comets, Meteors 1993* (A. Milani, M. Di Martino, A. Cellino, Eds., Kluwer Academic Publishers, Dordrecht), 271–296, 1994.
24. Y. G. Shkuratov, K. Muinonen, E. Bowell, K. Lumme, J. I. Peltoniemi, M. A. Kreslavsky, D. G. Stankevich, V. P. Tishkovetz, N. V. Opanasenko, L. Y. Melkumova. A critical review of theoretical models of negatively polarized light scattered by atmosphereless solar system bodies: *Earth, Moon, and Planets* **65**, 201–246, 1994.
25. K. Muinonen. Coherent backscattering of light by complex random media of spherical scatterers: Numerical solution: *Waves in Random Media* **14** 365–388, 2004.
26. H. Parviainen, K. Muinonen. Rough-surface shadowing for self-affine random rough surfaces: *J. Quantit. Spectrosc. Radiat. Transf.* **106**, 398–416, 2007.
27. H. Parviainen, K. Muinonen. Bidirectional reflectance of rough particulate media: ray-tracing solution: *J. Quantit. Spectrosc. Radiat. Transf.* **110**, 1418–1440, 2009.
28. J. I. Peltoniemi, K. Lumme. Light scattering by closely packed particulate media: *J. Opt. Soc. Am. A* **9**, 1320–1326, 1992.
29. K. Lumme, J. I. Peltoniemi, W. M. Irvine. Diffuse reflection in stochastically bounded semi-infinite media: *Trans. Theory Stat. Phys.* **19**, 317–332, 1990.
30. K. Lumme, J. Rahola. Comparison of light scattering by stochastically rough spheres, best-fit spheroids and spheres: *J. Quantit. Spectrosc. Radiat. Transf.* **60**, 439–450, 1998.
31. E. Zubko, Y. Shkuratov, K. Muinonen, G. Videen. Collective effects by agglomerated debris particles in the backscatter: *J. Quantit. Spectrosc. Radiat. Transf.* **100**(1-3), 489–495, 2006.
32. K. Muinonen, E. Zubko, J. Tyynelä, Y. G. Shkuratov, G. Videen. Light scattering by Gaussian random particles with discrete-dipole approximation: *J. Quantit. Spectrosc. Radiat. Transf.* **106**, 360–377, 2007.
33. J. Tyynelä, E. Zubko, G. Videen, K. Muinonen. Interrelating angular scattering characteristics to internal electric fields for wavelength-scale spherical particles: *J. Quantit. Spectrosc. Radiat. Transf.* **106**, 520–534, 2007.

34. J. Tyynelä, K. Muinonen, E. Zubko, G. Videen. Interrelating scattering characteristics to internal electric fields for Gaussian-random-sphere particles: *J. Quantit. Spectrosc. Radiat. Transf.* **109**, 2207–2218, 2008.
35. E. Zubko, K. Muinonen, Y. G. Shkuratov, G. Videen, T. Nousiainen. Scattering of light by roughened Gaussian random particles: *J. Quantit. Spectrosc. Radiat. Transf.* **106**, 604–615, 2007.
36. H. Lindqvist, K. Muinonen, T. Nousiainen. Light scattering by coated Gaussian and aggregate particles: *J. Quantit. Spectrosc. Radiat. Transf.* **110**, 1398–1410, 2009.
37. K. Muinonen, J. Tyynelä, E. Zubko, G. Videen, Y. G. Shkuratov. Single-scattering mechanism for negative polarization and opposition brightening of atmosphereless solar-system objects: *Bull. Am. Astron. Soc.* **39**, 429, 2007 (abstract).
38. Y. G. Shkuratov, A. Ovcharenko, E. Zubko, H. Volten, O. Muñoz, G. Videen. The negative polarization of light scattered from particulate surfaces and of independently scattering particles: *J. Quantit. Spectrosc. Radiat. Transf.* **88**, 267–284, 2004.
39. E. M. Purcell, and C. R. Pennypacker. Scattering and absorption of light by nonspherical dielectric grains: *Astrophys. J.* **186**, 705–714, 1973.
40. B. T. Draine. The discrete-dipole approximation and its application to interstellar graphite grains. *Astrophys. J.* **333**, 848–872, 1988.
41. A. G. Hoekstra, P. M. A. Sloot. Dipolar unit size in coupled-dipole calculations of the scattering matrix elements: *Opt. Lett.* **18**, 1211–1213, 1993.
42. K. Lumme, J. Rahola. Light scattering by porous dust particles in the discrete-dipole approximation: *Astrophys. J.* **425**, 653–667, 1994.
43. E. S. Zubko, Yu. G. Shkuratov, M. Hart, J. Eversole, G. Videen. Backscattering and negative polarization of agglomerate particles: *Opt. Lett.* **28**, 1504–1506, 2003.
44. K. Muinonen. Light scattering by stochastically shaped particles. In *Light Scattering by Nonspherical Particles: Theory, Measurements, and Applications* (M. I. Mishchenko, J. W. Hovenier, and L. D. Travis, Eds., Academic Press, San Diego), 323–352, 2000.
45. K. Muinonen, T. Nousiainen, P. Fast, K. Lumme, J. I. Peltoniemi. Light scattering by Gaussian random particles: ray optics approximation: *J. Quantit. Spectrosc. Radiat. Transf.* **55**, 577–601, 1996.
46. J. I. Peltoniemi, K. Lumme, K. Muinonen, W. M. Irvine. Scattering of light by stochastically rough particles: *Appl. Opt.* **28**, 4088–4095, 1989.
47. K. Muinonen, L. Lamberg, P. Fast, K. Lumme. Ray optics regime for Gaussian random sphere: *J. Quantit. Spectrosc. Radiat. Transf.* **57**, 197–205, 1997.
48. H. Volten, O. Muñoz, J. F. de Haan, W. Vassen, J. Hovenier, K. Muinonen, T. Nousiainen. Scattering matrices of mineral aerosol particles at 441.6 nm and 632.8 nm: *J. Geophys. Res.* **106**, 17375–17401, 2001.
49. T. Nousiainen, K. Muinonen, P. Räisänen. Scattering of light by large Saharan dust particles in a modified ray-optics approximation. *J. Geophys. Res. (Atmospheres)* **108**, D1, 4025, doi:10.1029/2001JD001277, 2003.
50. O. Muñoz, H. Volten, J. W. Hovenier, T. Nousiainen, K. Muinonen, D. Guirado, F. Moreno, L. B. F. M. Waters. Scattering matrix of large Saharan dust particles: experiments and computations. *Journal of Geophysical Research (Atmospheres)* **112**, D13215, doi:10.1029/2006JD008074, 2007.
51. K. Muinonen, T. Nousiainen, H. Lindqvist, O. Muñoz, G. Videen. Light scattering by Gaussian particles with internal inclusions and roughened surfaces using ray optics: *J. Quantit. Spectrosc. Radiat. Transf.* **110**, 1628–1639, 2009.
52. K. Muinonen. Light scattering by Gaussian random particles: Rayleigh and Rayleigh-Gans approximations: *J. Quantit. Spectrosc. Radiat. Transf.* **55**, 603–613, 1996.

53. A. Battaglia, K. Muinonen, T. Nousiainen, and J. I. Peltoniemi. Light scattering by Gaussian particles: Rayleigh-ellipsoid approximation: *J. Quantit. Spectrosc. Radiat. Transf.* **63**, 277–303, 1999.
54. T. Nousiainen, K. Muinonen, J. Avelin, A. Sihvola. Microwave backscattering by nonspherical ice particles at 5.6 GHz using second-order perturbation series: *J. Quantit. Spectrosc. Radiat. Transf.* **70**: 639–661 2001.
55. B. Veihelmann, T. Nousiainen, M. Kahnert, W. J. van der Zande. Light scattering by small feldspar particles simulated using the Gaussian random sphere geometry: *J. Quantit. Spectrosc. Radiat. Transf.* **100**, 393–405 2006.
56. T. Nousiainen, K. Muinonen. Surface-roughness effects on single-scattering properties of wavelength-scale particles. *J. Quantit. Spectrosc. Radiat. Transf.* **106**, 389–397, 2007.
57. W. Sun, T. Nousiainen, K. Muinonen, Q. Fu, N. G. Loeb, G. Videen. Light scattering by Gaussian particles: a solution with finite-difference time domain technique: *J. Quantit. Spectrosc. Radiat. Transf.* **79-80**, 1083–1090, 2003.
58. H. C. van de Hulst: *Light Scattering by Small Particles*, Wiley, New York, 1957.
59. E. Zubko, Y. G. Shkuratov, G. Videen, K. Muinonen. Effects of interference on the backscattering properties of irregularly shaped particles using DDA. In *9th Conference on Electromagnetic and Light Scattering by Nonspherical Particles: Theory, Measurements, and Applications, Book of Abstracts* (N. Voshchinnikov, Ed., St. Petersburg, Russia, June 5-9, 2006), 287–290, 2006.
60. K. Muinonen, J. Tyynelä, E. Zubko, G. Videen. Scattering parameterization for interpreting asteroid polarimetric and photometric phase effects: *Earth, Planets, and Space*, in press, 2009.
61. J. E. Hansen, L. D. Travis. Light scattering in planetary atmospheres: *Space Sci. Rev.* **16**, 527–610, 1974.
62. Y. Kuga, A. Ishimaru. Retroreflectance from a dense distribution of spherical particles: *J. Opt. Soc. Am. A* **1**, 831–835, 1984.
63. L. Tsang, A. Ishimaru. Backscattering enhancement of random discrete scatterers: *J. Opt. Soc. Am. A* **1**, 836–839 1984.
64. L. Tsang, A. Ishimaru. Theory of backscattering enhancement of random discrete isotropic scatterers based on the summation of all ladder and cyclical terms: *J. Opt. Soc. Am. A* **2**, 1331–1338, 1985.
65. P.-E. Wolf, G. Maret. Weak localization and coherent backscattering of photons in disordered media: *Phys. Rev. Letters* **55**, 2696–2699, 1985.
66. K. Muinonen. Coherent backscattering of light by solar system bodies: Efficient scalar computation: *Bull. Am. Astron. Soc.* **31**, 1076, 1999 (abstract).
67. S. Kaasalainen, K. Muinonen, J. Piironen. Comparative study on opposition effect of icy solar system objects. *J. Quantit. Spectrosc. Radiat. Transf.* **70**, 529–543, 2001.
68. K. Muinonen, H. Parviainen, J. Näränen, J.-L. Josset, S. Beauvivre, P. Pinet, S. Chevrel, D. Koschny, B. Grieger, B. Foing. Lunar single-scattering, porosity, and surface roughness properties with SMART-1/AMIE. *Astronomy & Astrophysics*, submitted, 2009.
69. M. I. Mishchenko, J. M. Dlugach. Accuracy of the scalar approximation in computations of diffuse and coherent backscattering by discrete random media. *Phys. Rev. A* **78**, 063822, 2008.
70. M. I. Mishchenko J.-M. Luck, T. M. Nieuwenhuizen. Full angular profile of the coherent polarization opposition effect: *J. Opt. Soc. Am. A* **17**, 888–891, 2000.
71. J. M. A. Danby: *Fundamentals of Celestial Mechanics*, Willmann-Bell, Inc., Richmond, VA, 1992.

72. W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery: *Numerical Recipes in Fortran, The Art of Scientific Computing, 2nd edn*, Cambridge University Press, 1992.
73. B. Zellner, T. Gehrels, J. Gradie. Minor planets and related objects. XVI. Polarimetric diameters: *Astron. J.* **79**, 1100–1110, 1974.
74. B. Zellner, J. Gradie. Minor planets and related objects. XX. Polarimetric evidence for the albedos and compositions of 94 asteroids: *Astron. J.* **81**, 262–280, 1976.
75. G. Videen, K. Muinonen, K. Lumme. Coherence, power laws, and the negative polarization surge: *Appl. Opt.* **42**, 3647–3652, 2003.
76. K. Lumme, J. Rahola, and J. W. Hovenier. Light scattering by dense clusters of spheres: *Icarus* **126**, 455–469, 1997.
77. M. I. Mishchenko, J. M. Dlugach, L. Liu, V. K. Rosenbush, N. N. Kiselev, and Y. G. Shkuratov. Direct solutions of the maxwell equations explain opposition phenomena observed for high-albedo solar system objects: *Astrophys. J.* **705**, L118–L122, 2009.

Color Section

Chapter 3

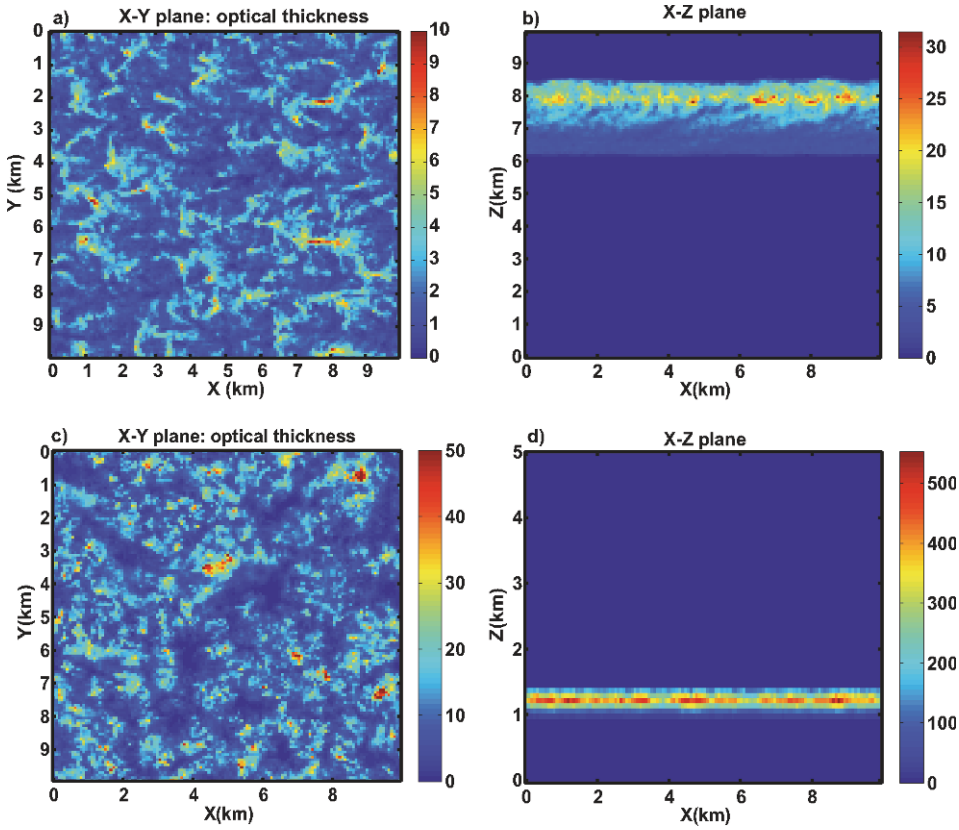


Fig. 3.8. Synthetic clouds used for the radiative transfer. (a) and (b) cirrus cloud with a mean optical thickness of 2. (c) and (d) stratocumulus cloud with a mean optical thickness of 10. Optical thickness integrated along the z-axis (a) and (c), and along the y-axis (b) and (d). (a) and (b) are from Cornet et al. (2009).

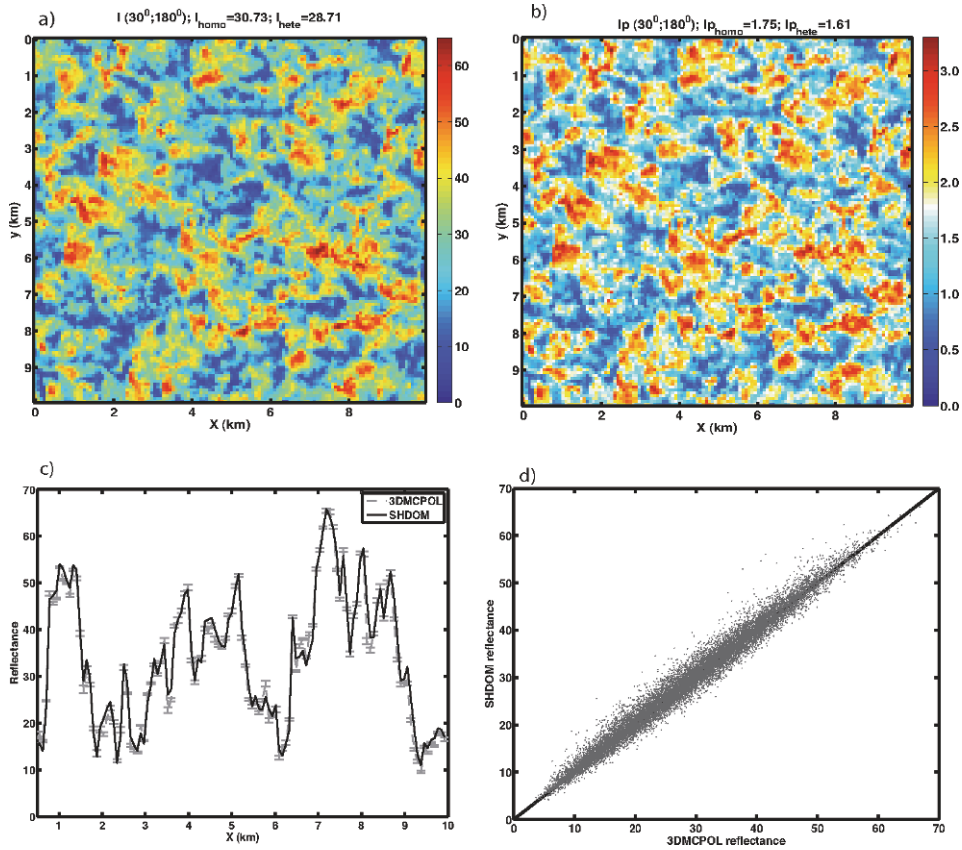


Fig. 3.9. Total (a) and polarized (b) reflectances at 78 m resolution for $\theta_s = 60°$; $\theta_v = 30$ and $\varphi_v = 180°$ (forward direction) for the cirrus cloud presented in Fig. (3.8). The wavelength is 865 nm. The mean values are indicated above the figures for the 3-D cirrus cloud (hete) and for the homogeneous cloud (homo). (c) and (d) comparisons with SHDOM for the total reflectance value: (c) horizontal cut along the line 75 corresponding to about 6 km and (d) scatterplot between SHDOM and 3DMCPOL reflectances for the entire field presented in (a). Error bars correspond to the statistical errors of Monte–Carlo simulations. (a) and (b) are from Cornet et al. (2009).

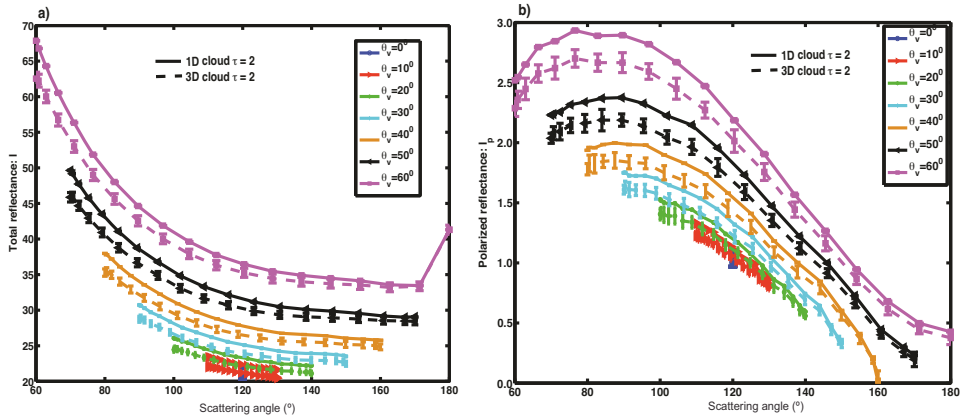


Fig. 3.10. Total (a) and polarized (b) reflectances for $\theta_s = 60^\circ$ as a function of the scattering angles ($\Theta = 180^\circ$ corresponding to backward scattering). Solid lines are for a homogeneous cirrus cloud of optical thickness 2 and dashed lines for the spatial mean reflectance values for the cirrus cloud presented in Fig. 3.8. Error bars corresponds to the statistical error of Monte–Carlo computations. Modified from Cornet et al. (2009).

Chapter 6

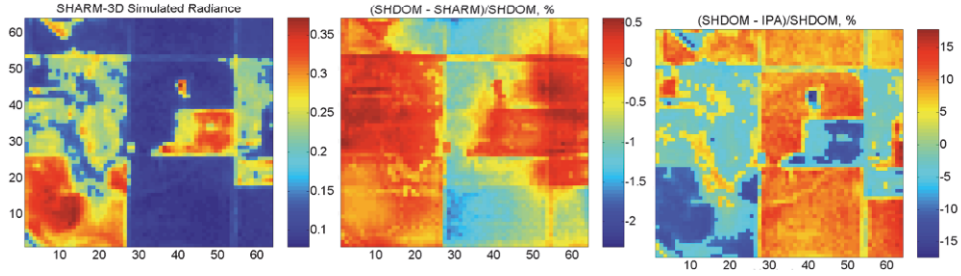


Fig. 6.4. Comparison of code SHARM-3D with SHDOM and independent pixel approximation (IPA) (from Lyapustin, 2002). The left image shows SHARM-3D simulated TOA radiance. The 64×64 pixels surface area at 30-m resolution was modeled using Landsat-7 reflectance in band 4 over the Oklahoma site on April 4, 2000. The middle and right images show the relative difference of SHARM-3D and IPA solutions with SHDOM.

Chapter 8

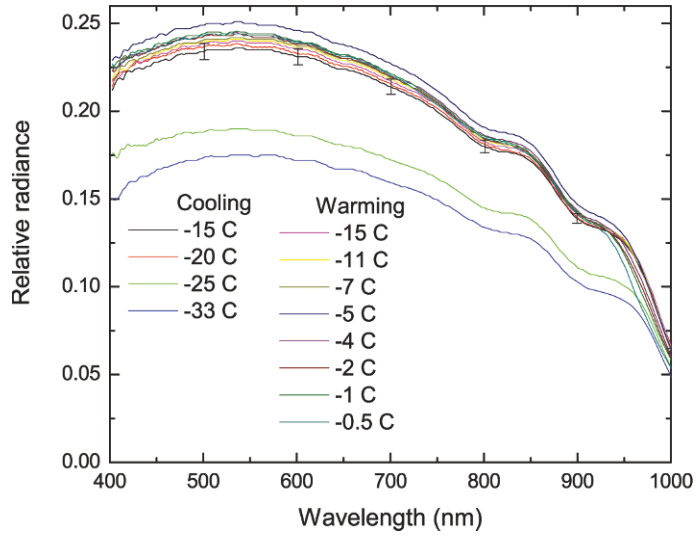


Fig. 8.17. Temperature-dependent changes in spectral radiance transmitted at 30° from nadir. Radiances were measured through the glass plate used to support the sample, then normalized by the incident irradiance.

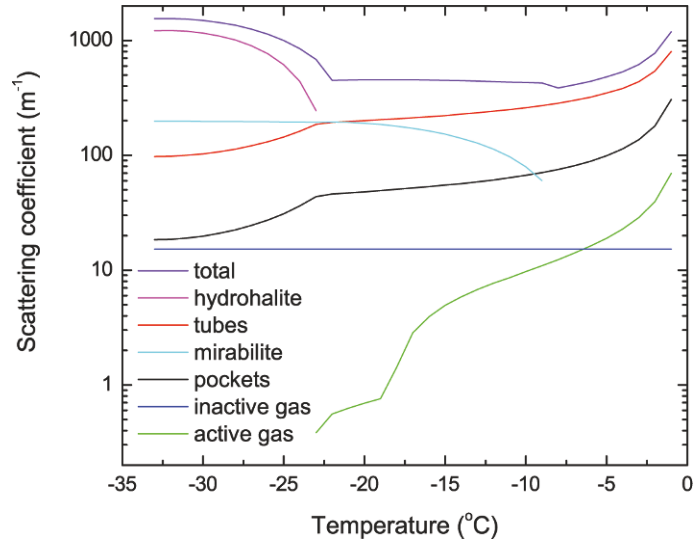


Fig. 8.21. Temperature-dependent scattering coefficients for various types of scatterers in sea ice as predicted by the structural-optical model.

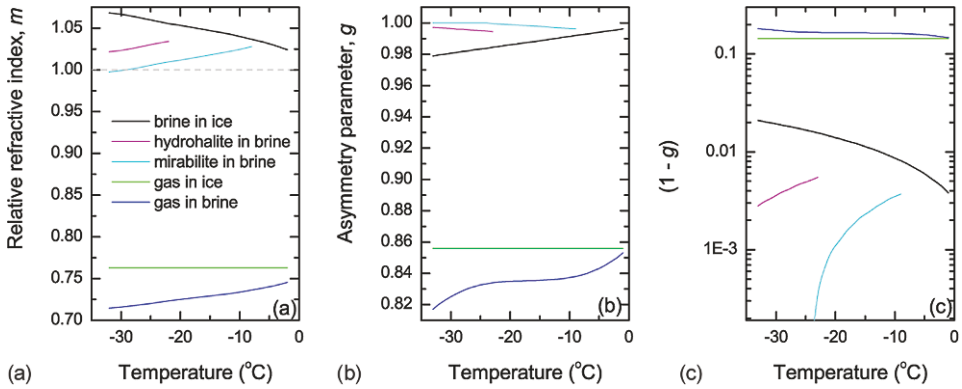


Fig. 8.22. Temperature-dependence of (a) relative refractive index, (b) asymmetry parameter, g, and (c) $(1-g)$ for inclusions of brine in ice, gas in ice, gas in brine, mirabilite in brine, and hydrohalite in brine. Asymmetry parameters were calculated using equivalent spheres and Mie theory.

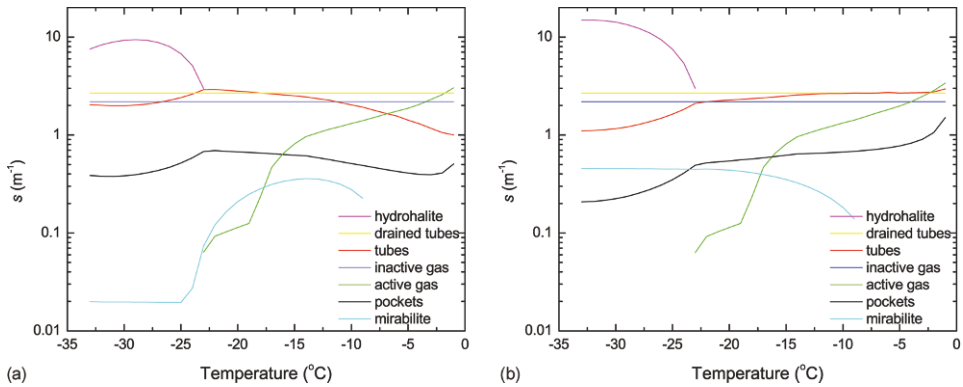


Fig. 8.27. Temperature-dependent partial similarity parameters for each constituent in the full structural-optical model using (a) $m_{\mathrm{brine}}(T)$ and (b) $m_{\mathrm{brine}}(T = -15°\mathrm{C})$.

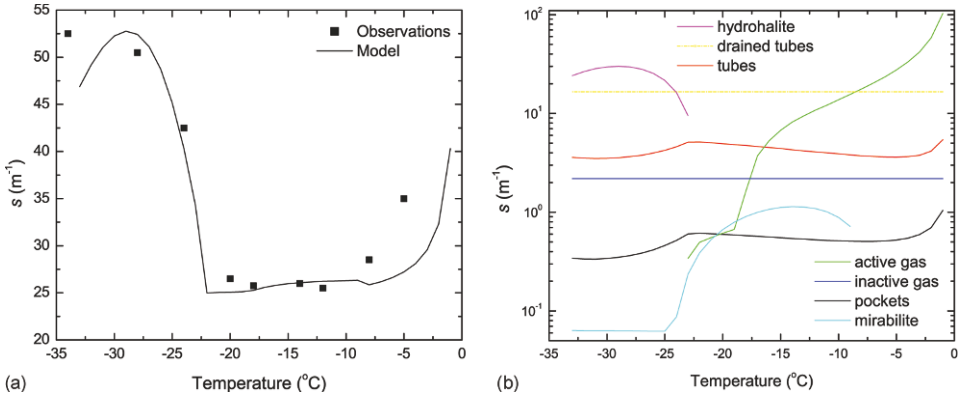


Fig. 8.28. Observations and model predictions for a high salinity sea ice sample, (a) comparison between s derived from observations and predicted with model, (b) model predicted similarity parameters for each constituent.

Chapter 9

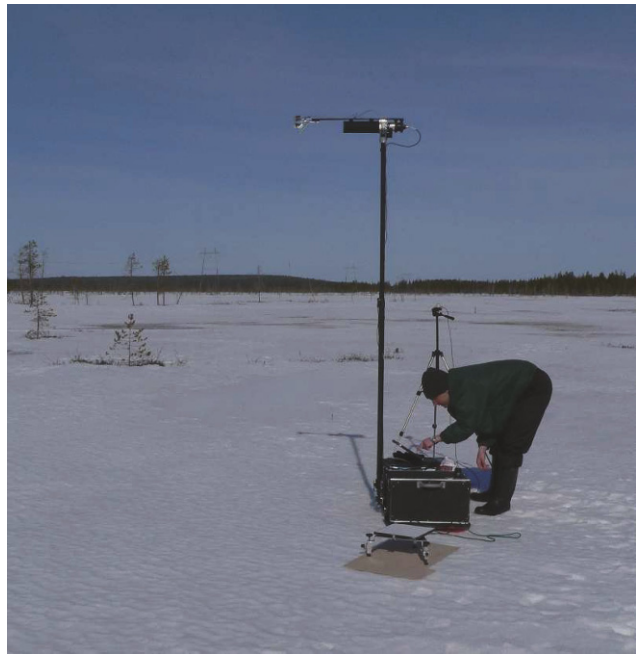


Fig. 9.3. The FIGIFIGO measuring BRF and polarization in Sodankylä in 2009. Optics is located horizontally in the upper end of the arm, and is looking down to the target through a mirror. The mirror has two small actuators to fine tune the spot location or scan over a small range. In front is the Spectralon white reference target, and behind the researcher's head is the sun and sky light monitoring pyranometer.

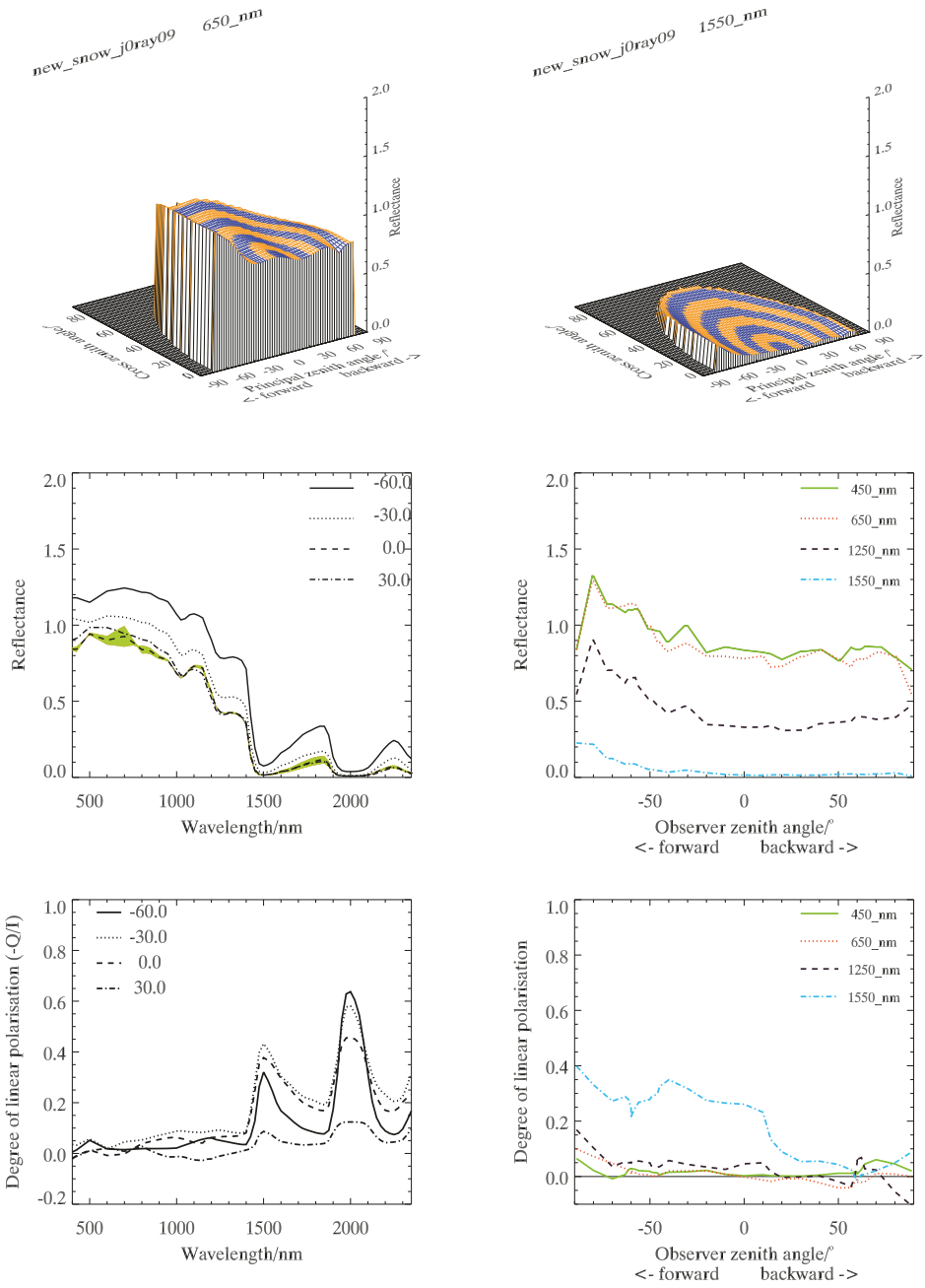


Fig. 9.6. Simulated BRFs of new snow, from top left: BRF cake in 650 and 1550 nm, the reflection spectra in four directions on the principal plane (colour indicates difference between two sampling), the BRF on the principal plane in four channels, the polarization spectra in four directions on the principal plane, and the polarization curve on the principal plane in four channels. The angle of incidence is 60°. Minor oscillations in the data are Monte Carlo noise.

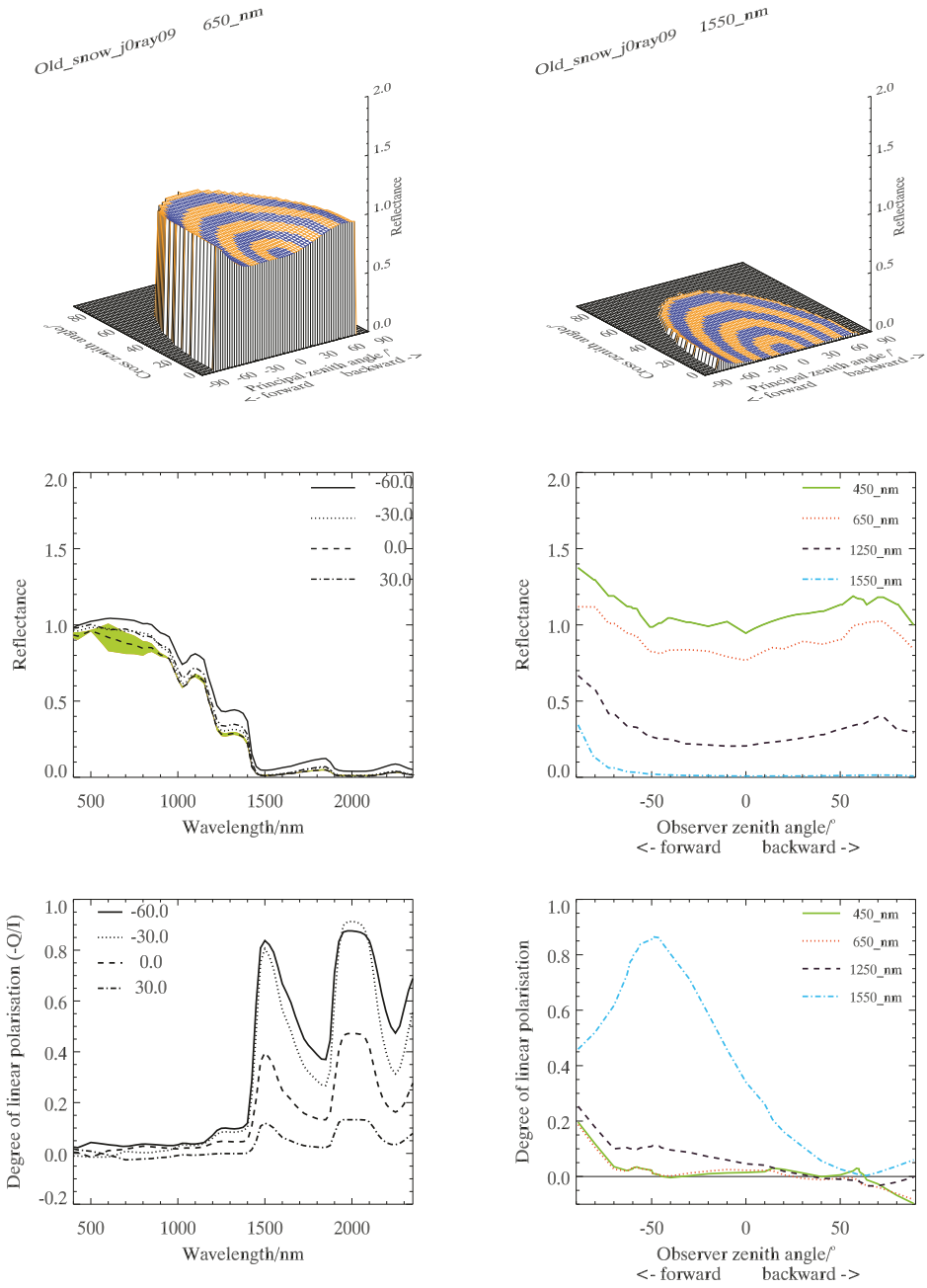


Fig. 9.7. Simulated BRFs of dry snow, as in Fig. 9.6. Note that the old snow spectrum is darker in infrared than in new snow, and polarization gets stronger. BRF shape changes are small.

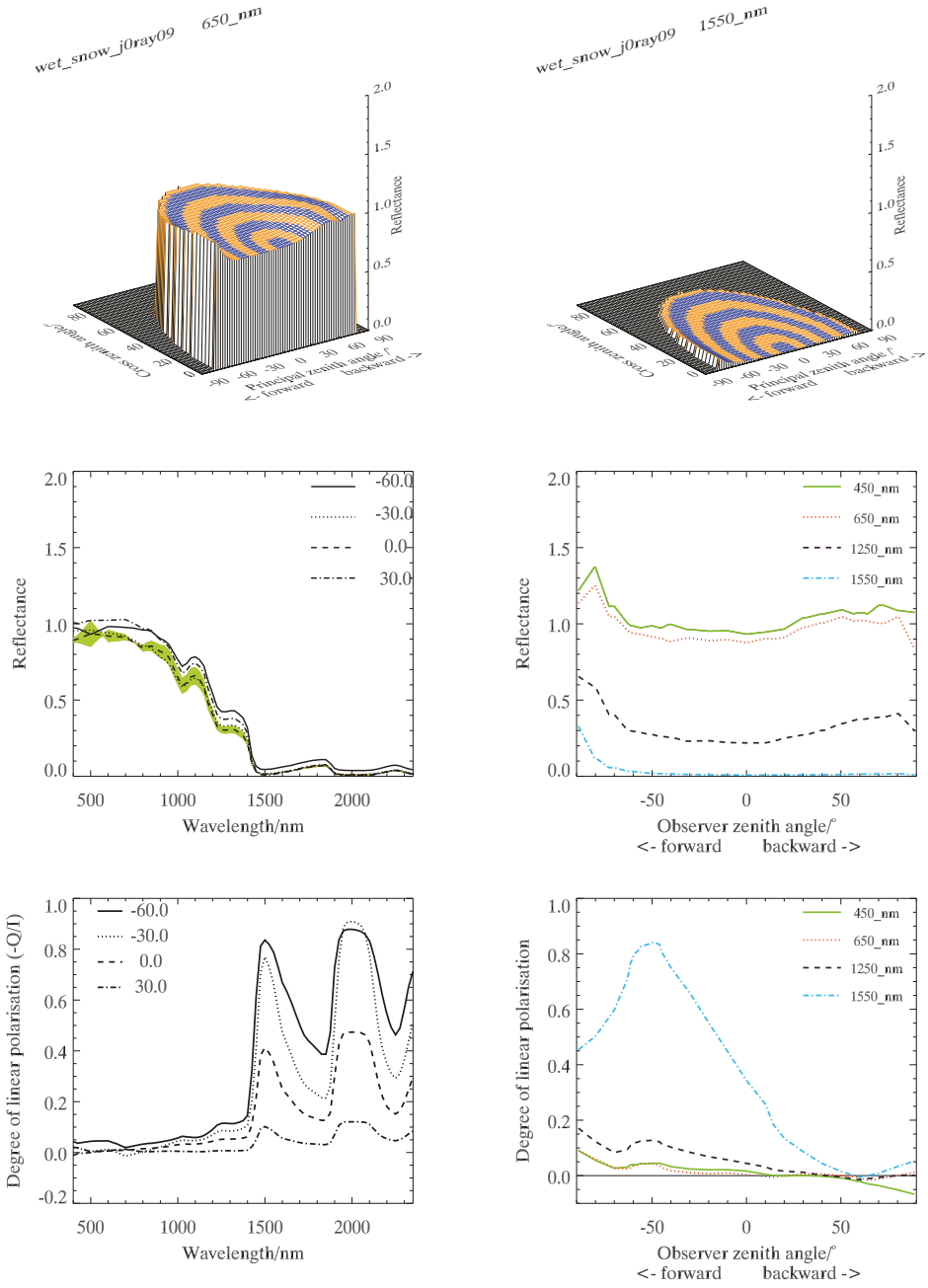
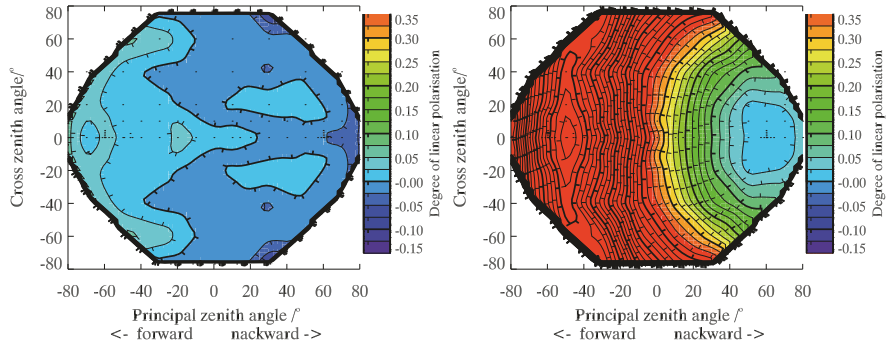
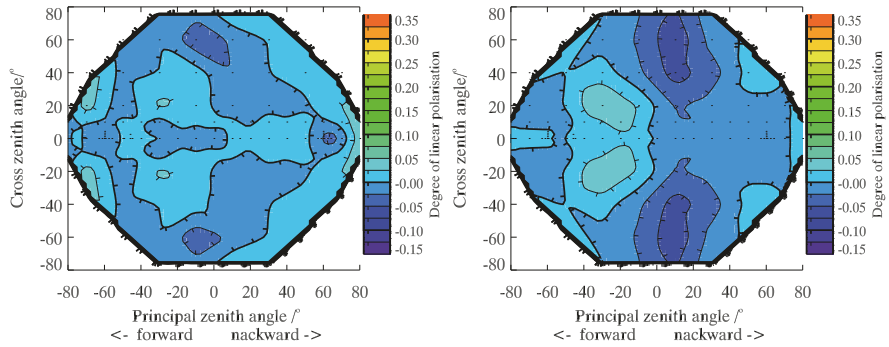


Fig. 9.8. Simulated BRFs of wet snow, as in Fig. 9.6. The only significant difference from dry snow is the very small spectral changes in some wavelengths, to be discussed in the next subsections.

$-Q/I$



$-U/I$



$-V/I$

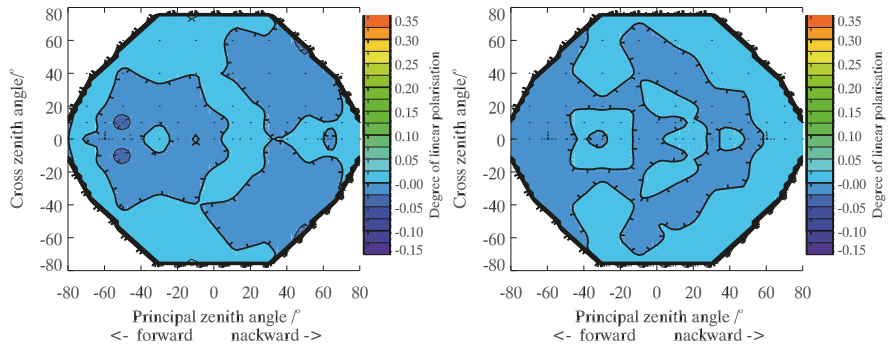


Fig. 9.9. Flat polarization plots of simulated dry snow, on the left 550 nm, on the right 1550 nm. From top to bottom $-Q/I$, $-U/I$, and $-V/I$. Polarisation is here defined in the scattering plane.

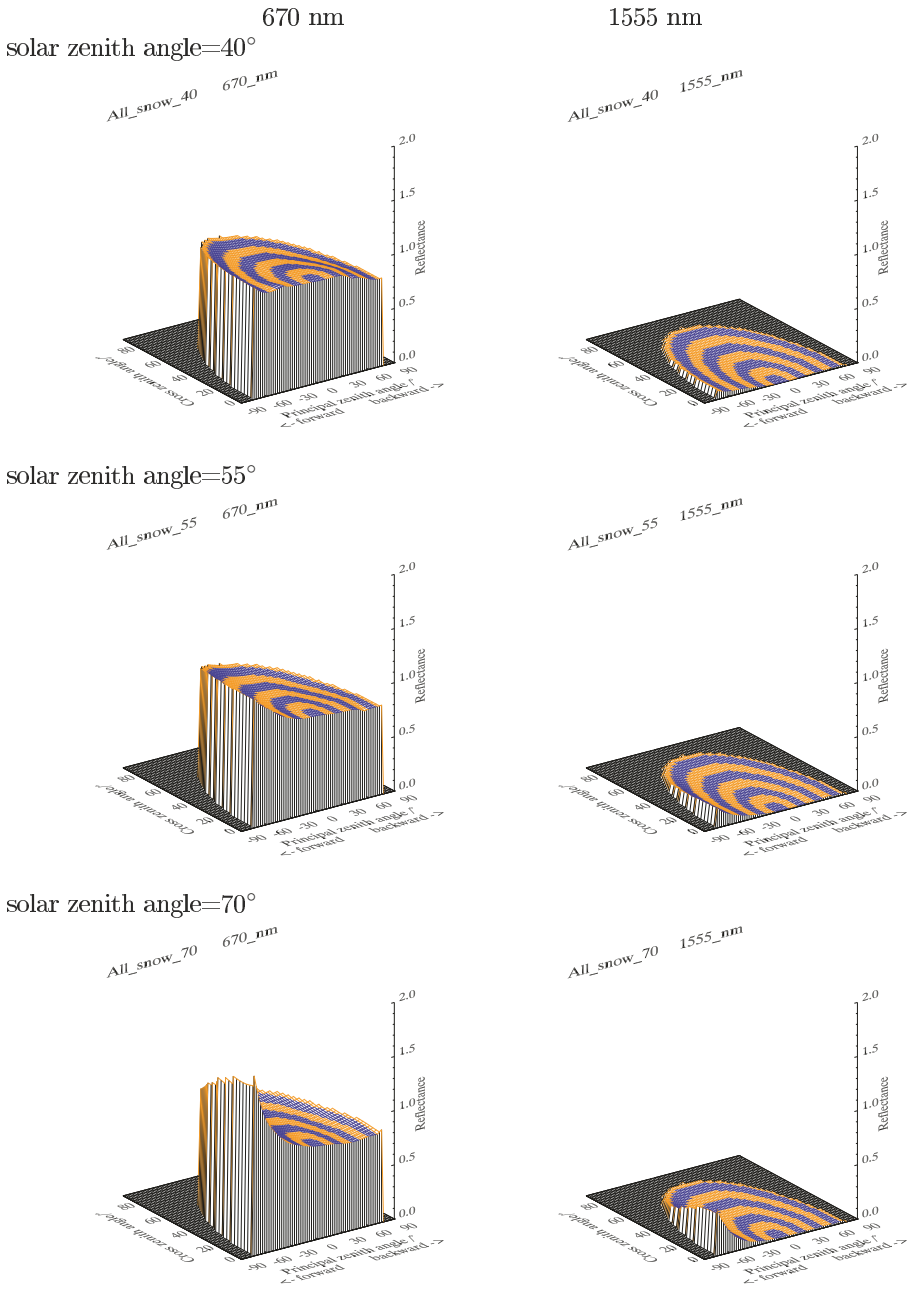


Fig. 9.10. The general 3-D shape of the snow BRF in varying angles of incidence, from top to bottom: 40°, 55°, 70°) and two wavelengths left 670 nm, right 1555 nm. The results are averaged over all measured samples, and show a typical strong forward scattering pattern, stronger in infrared than in visual, and stronger in larger zenith angles than smaller ones. Minor oscillations in the data are mostly caused by the small footprint size vs. surface roughness, and to some extent an averaging artifact.

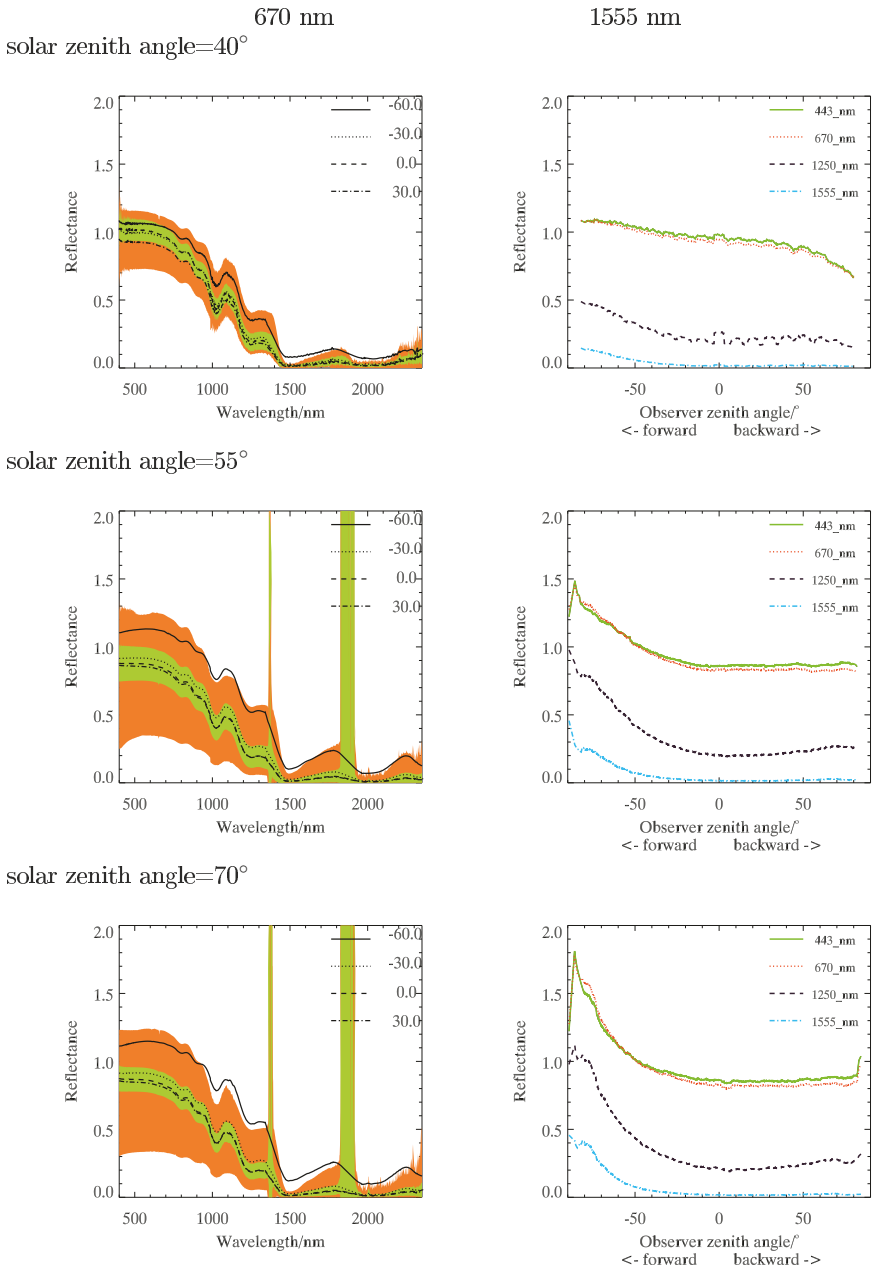


Fig. 9.11. Continuation of Fig. 9.10 with more details. On the left is the reflectivity spectra at four zenith angles on the principal plane (60° forward, 30° forward, nadir, and 30° backward. On the right is the BRF curve on the principal plane as a function of the zenith angle in four channels (violet 445 nm, red 670 nm, SWIR 1225 nm, SWIR 1555 nm). Green colour gives the standard deviation of all measured values at nadir and orange min/max values. The strong peaks and oscillations in some parts of the spectra are sensor noise, because the atmospheric water vapor absorbs all the incident signal in those wavelengths.

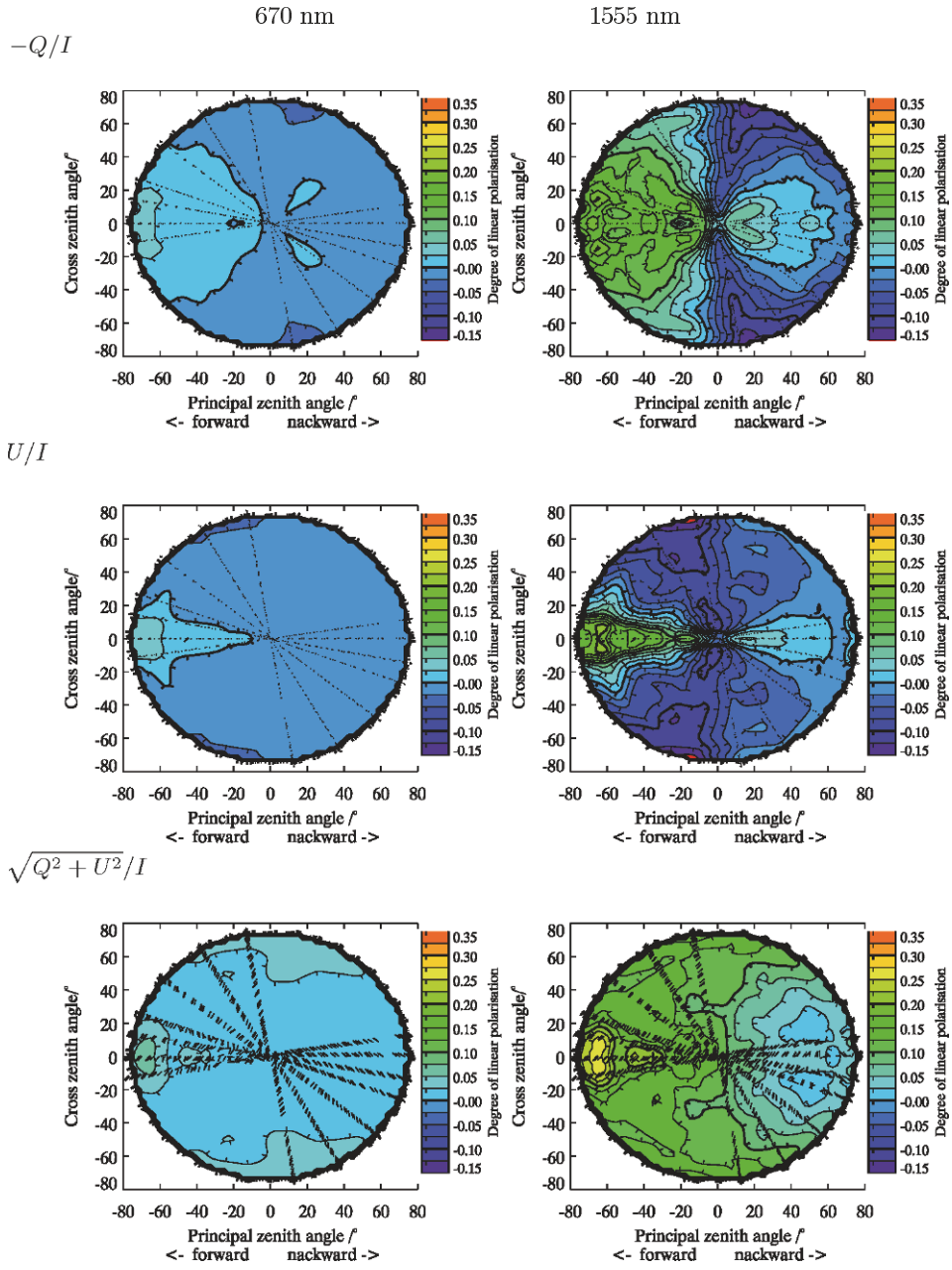


Fig. 9.13. Polarization properties of smooth snow from Mantovaaranaapa 2009. At the top the degree of linear horizontal–vertical polarization (Stokes $-Q/I$), in the middle U/I, and, at the bottom, total linear polarization $\sqrt{Q^2+U^2}/I$, taken at wavelengths of 670 nm (left) and 1555 nm (right), with an angle of incidence of 60°. The short lines in the bottom diagrams indicate the direction of polarization.

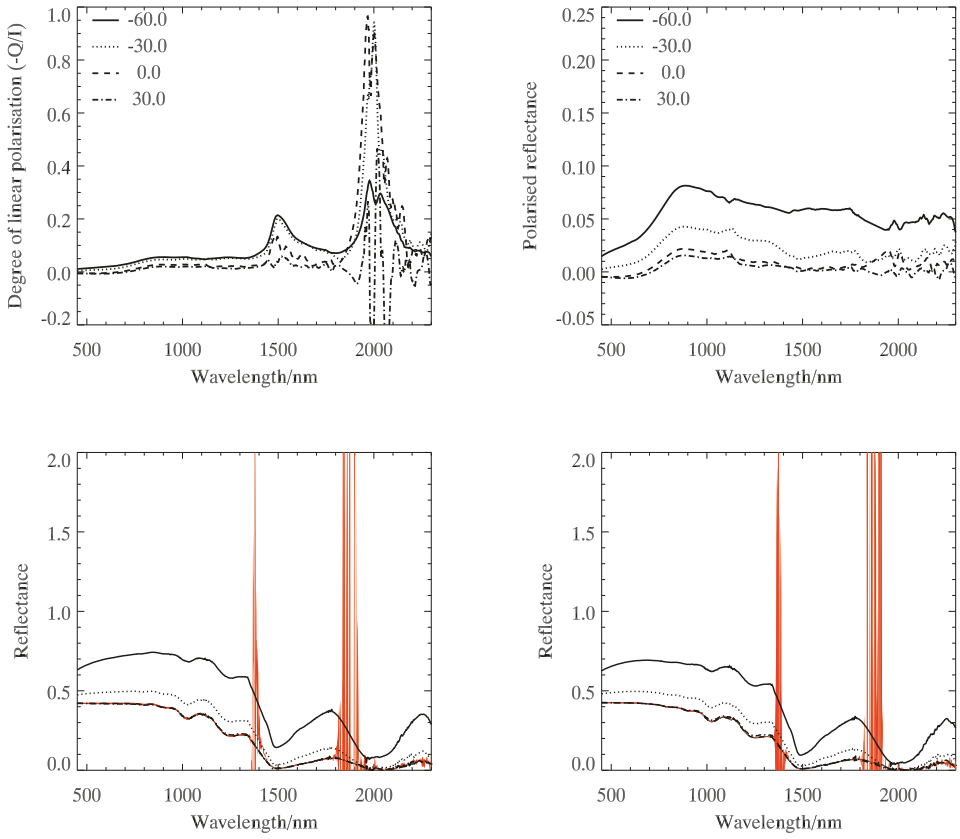


Fig. 9.14. The polarization spectrum shown in different ways. At the top left the degree of linear polarization (−Q/I), at right the polarized reflectance (−Q), and at the bottom horizontally and vertically polarized reflectances separately. All of these show interesting spectral features.

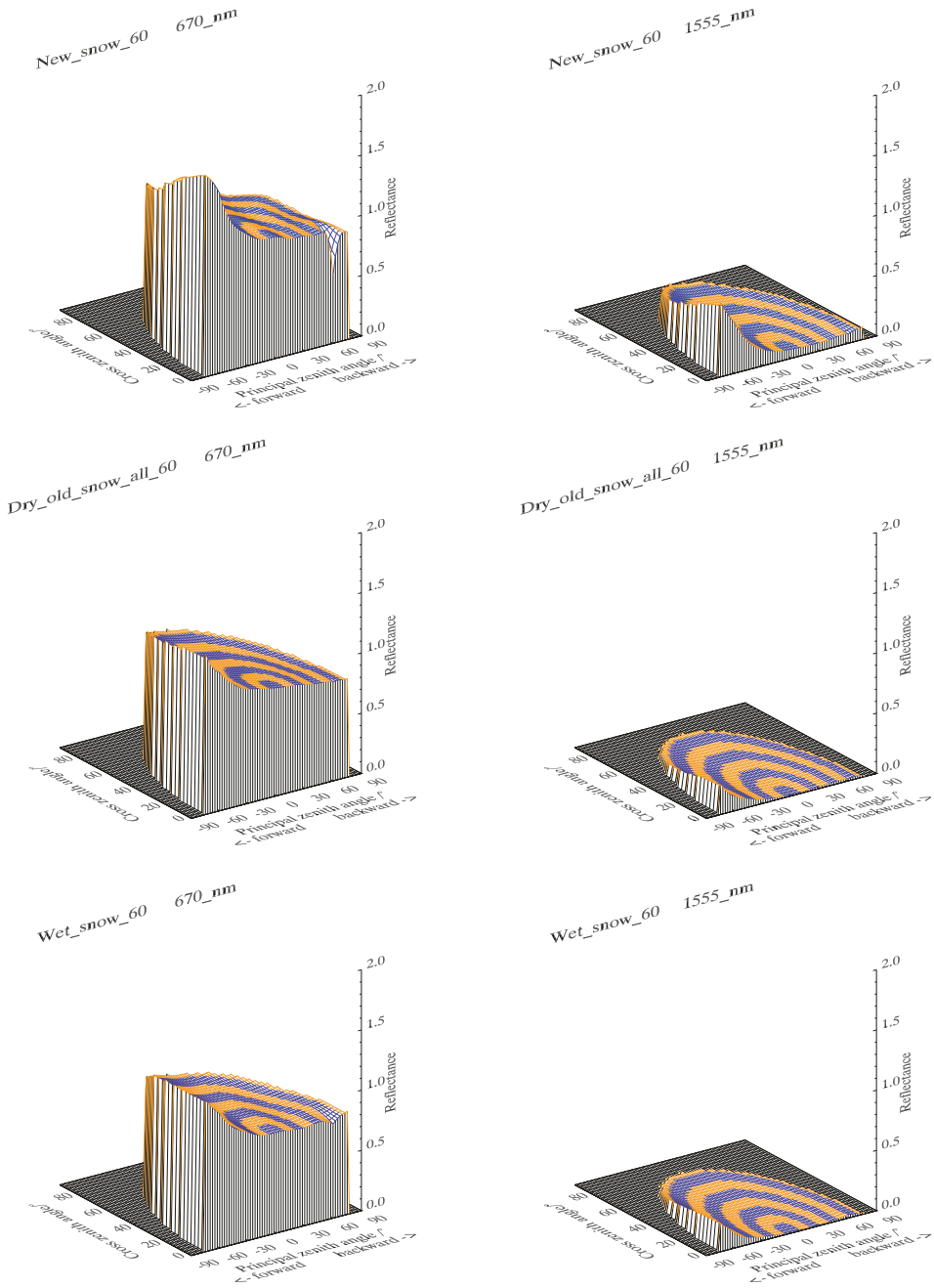


Fig. 9.15. The differences over the main snow types, new snow, old dry snow, and melting snow (from top to bottom), averaged over all samples in the class, otherwise as in Fig. 9.10, with an angle of incidence 60°.

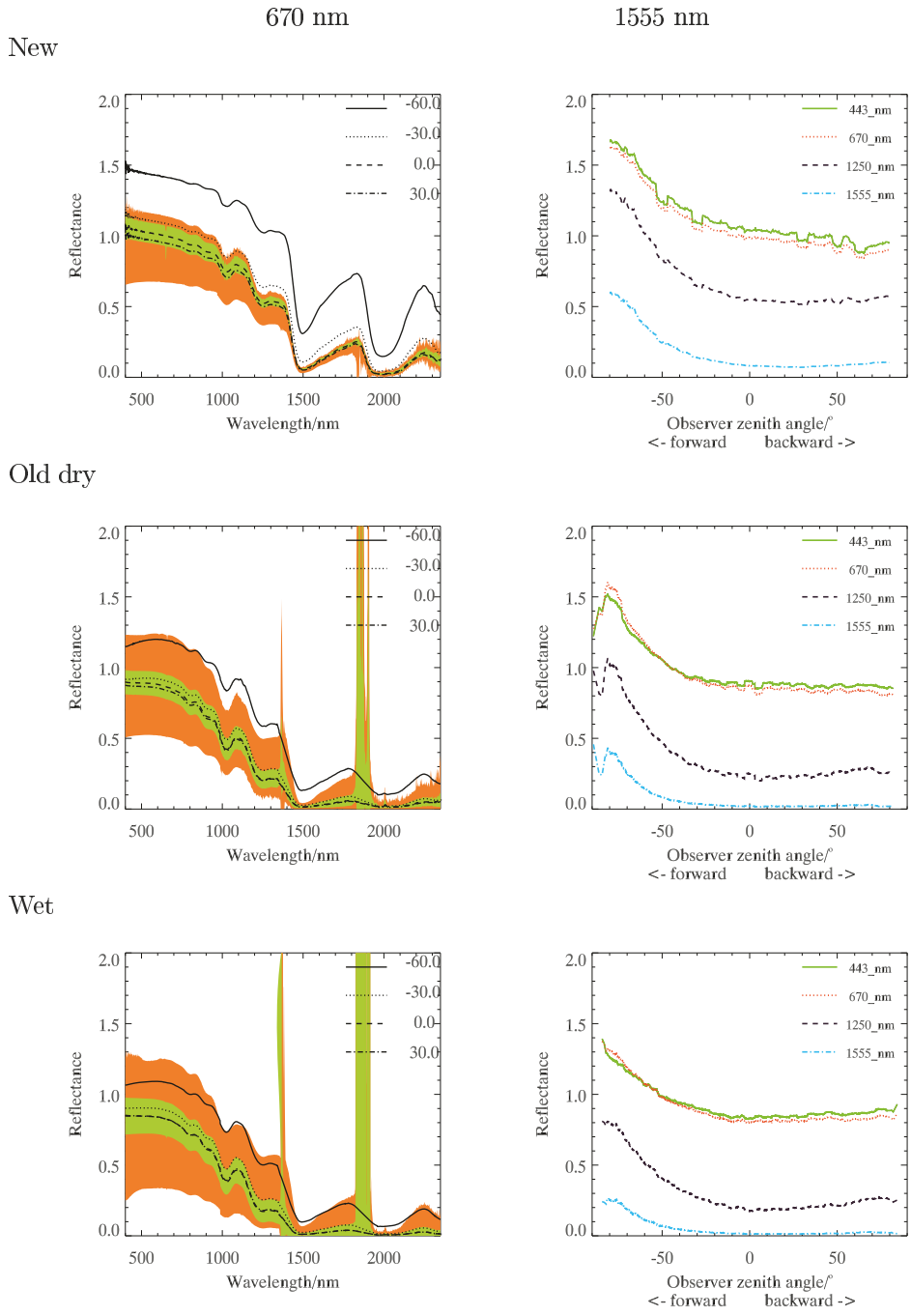
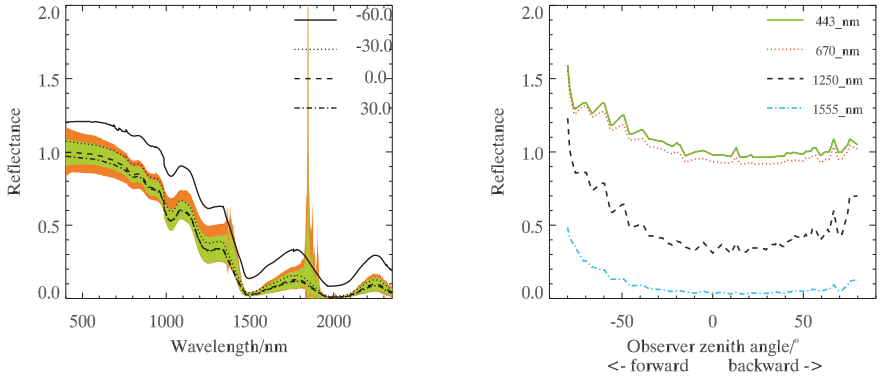
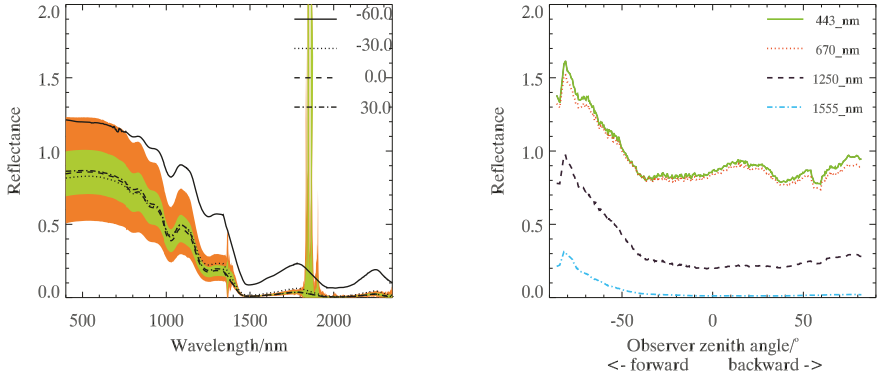


Fig. 9.16. Continuation of Fig. 9.15, analogous to Fig. 9.11. Here one can see more clearly differences in forward scattering and size effect in spectrum.

2005



2007



2008

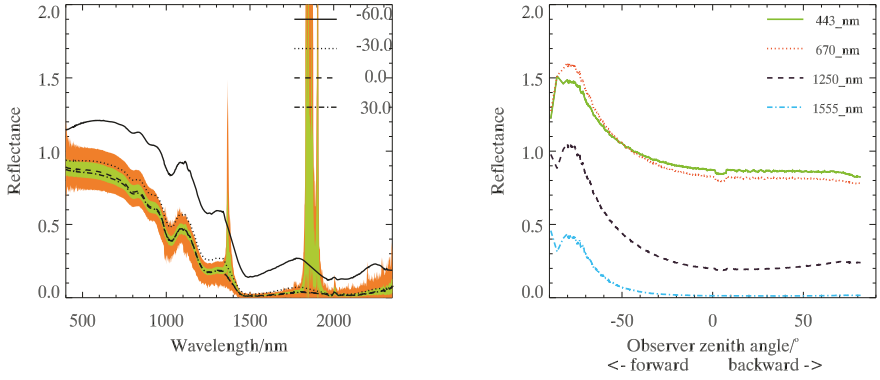
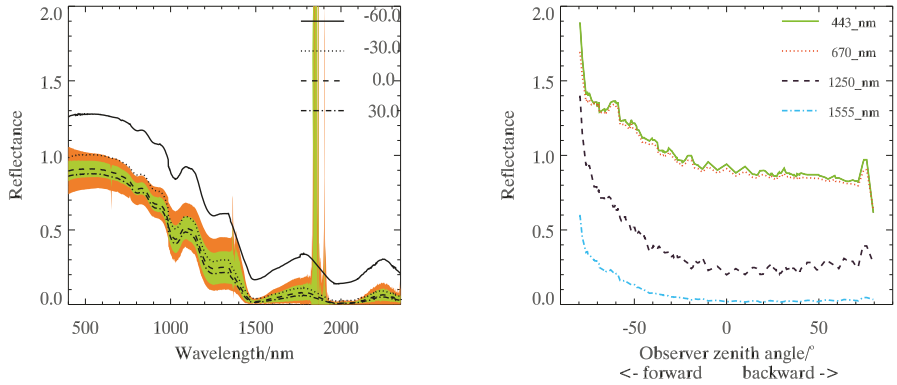
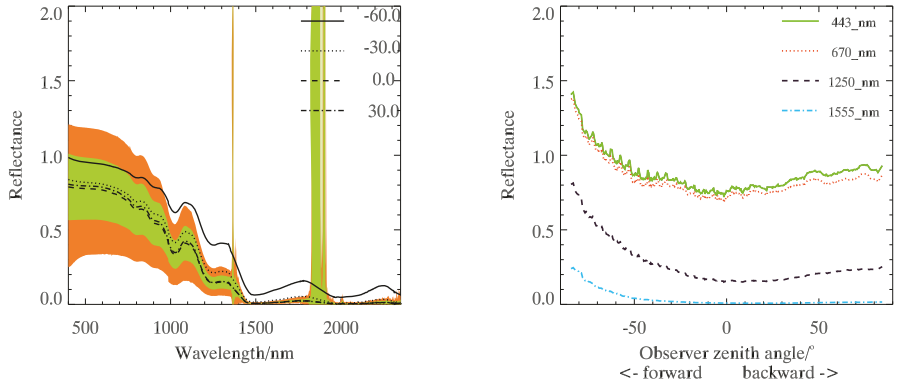


Fig. 9.17. Various dry old snow types, from top to bottom, 2005, 2007, 2008, as in Fig. 9.11, and with an angle of incidence 60°. From this figure one can already see, that samples differ, and single dry snow class is a clear oversimplification, but to keep this review in schedule, further dry snow analysis is dropped.

2005



2007



2009

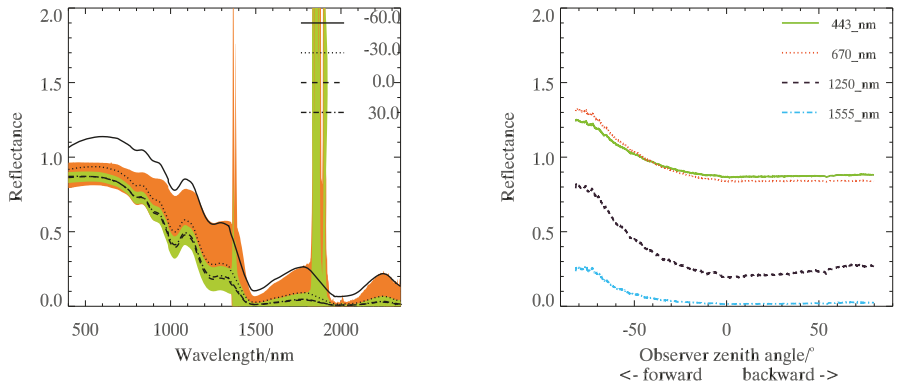


Fig. 9.18. Various wet and melting snow types, from top to bottom 2005, 2007 and 2009 snows, as in Fig. 9.11, and with an angle of incidence 60°. Note especially the small bump and drop at 60° to 70° forward.

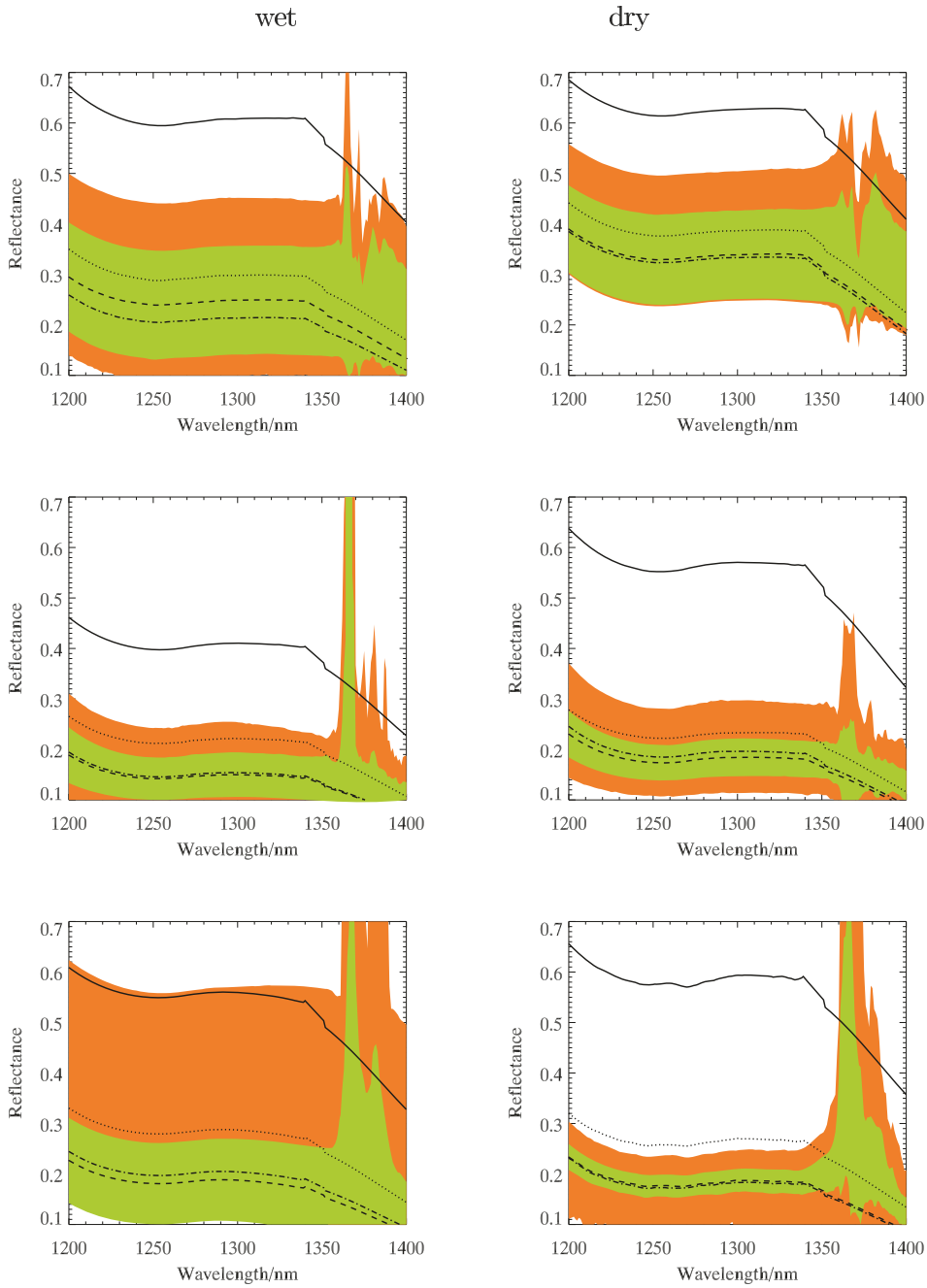


Fig. 9.19. More detailed view of the reflection spectra in the most wetness sensitive region, on the left wet snow, on the right dry old snow. The most indicative feature is the relative difference between 1250 and 1350 nm.

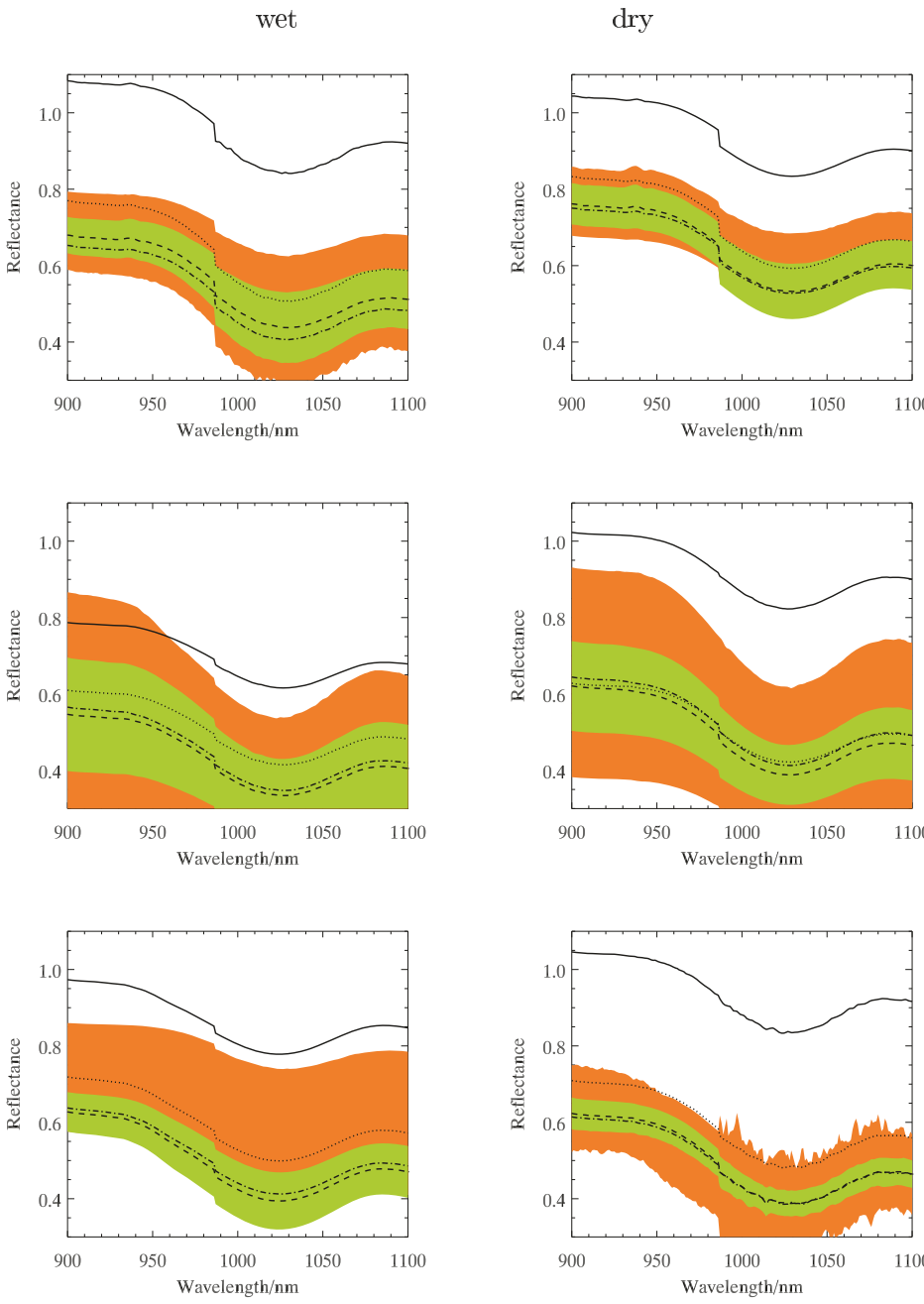


Fig. 9.20. Fig. 9.19 continues at spectral range 900 to 1100 nm. The most indicative feature here should be the difference between 960 nm and 1000 nm or the location of the minimum, although it is very difficult to see even from a zoomed curve, and numerical comparison is needed. Warning that the sensor border is here at 981 nm, which typically causes an erratic step of 1% to 5%, confusing analysis.

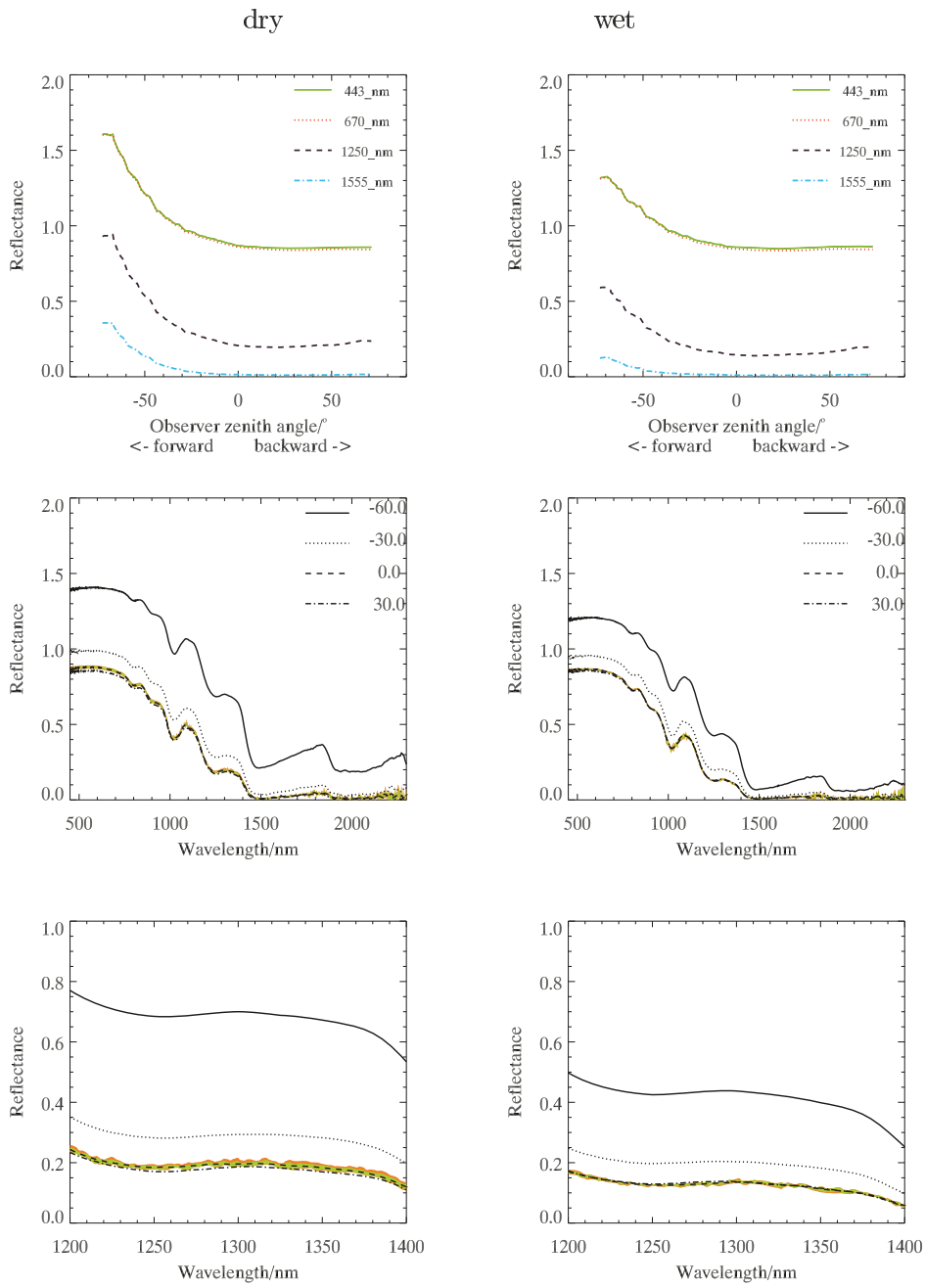


Fig. 9.21. Monitoring the melting process in laboratory. The same snow when dry (left) and after about 1 hour melting inside (right). At the top there is the BRF on the principal plane, in the middle the spectral plot from four directions in the principal plane, and at the bottom a zoom into the more wetness-sensitive spectral band. The light zenith angle was 60°.

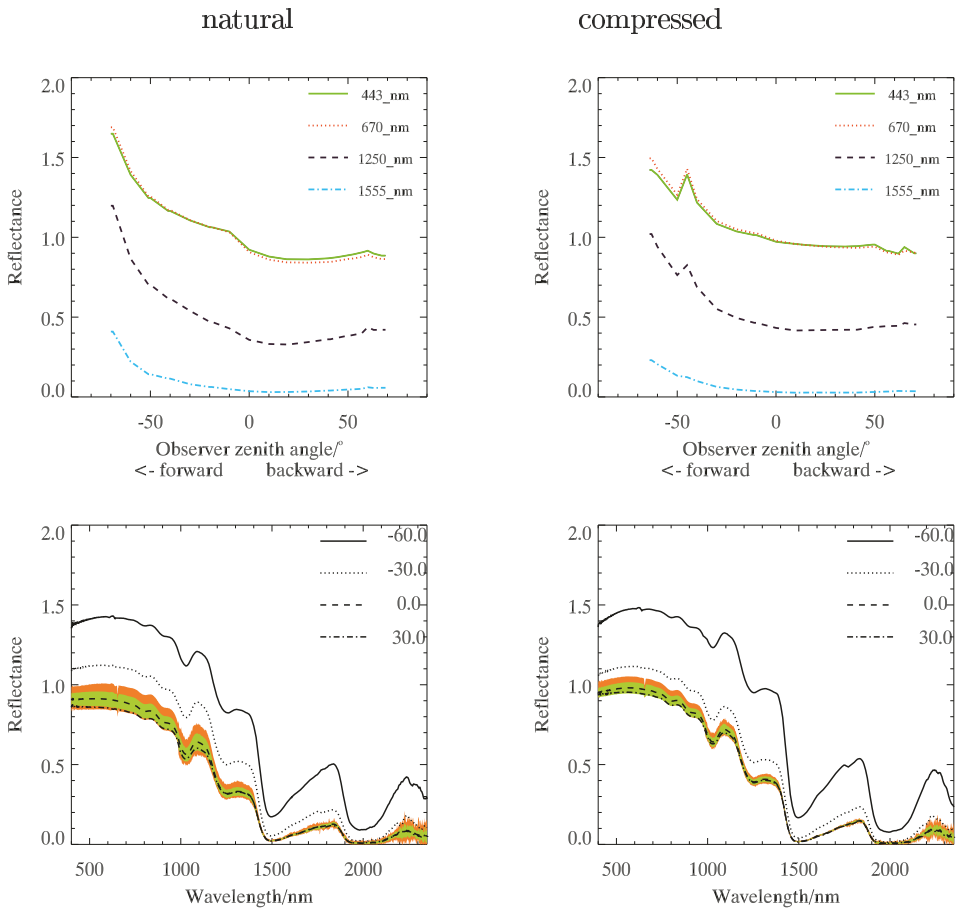


Fig. 9.23. Difference between natural (left) and compressed (right) snow.

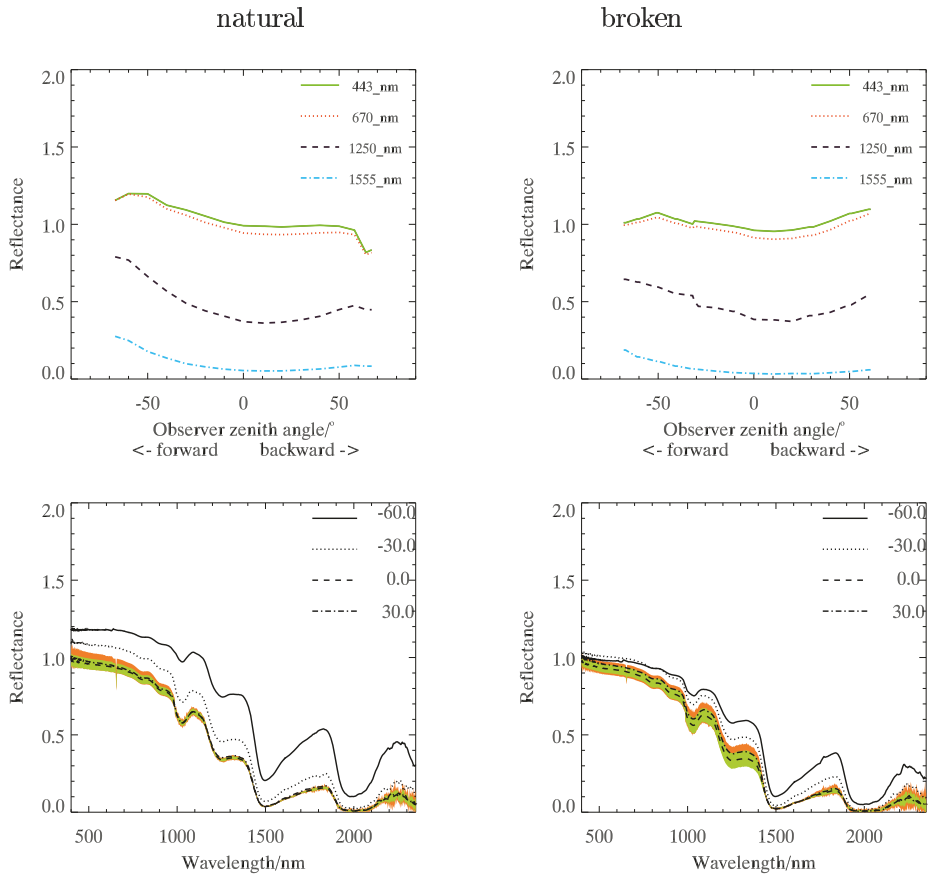


Fig. 9.24. Difference between natural (left) and broken (right) snow.

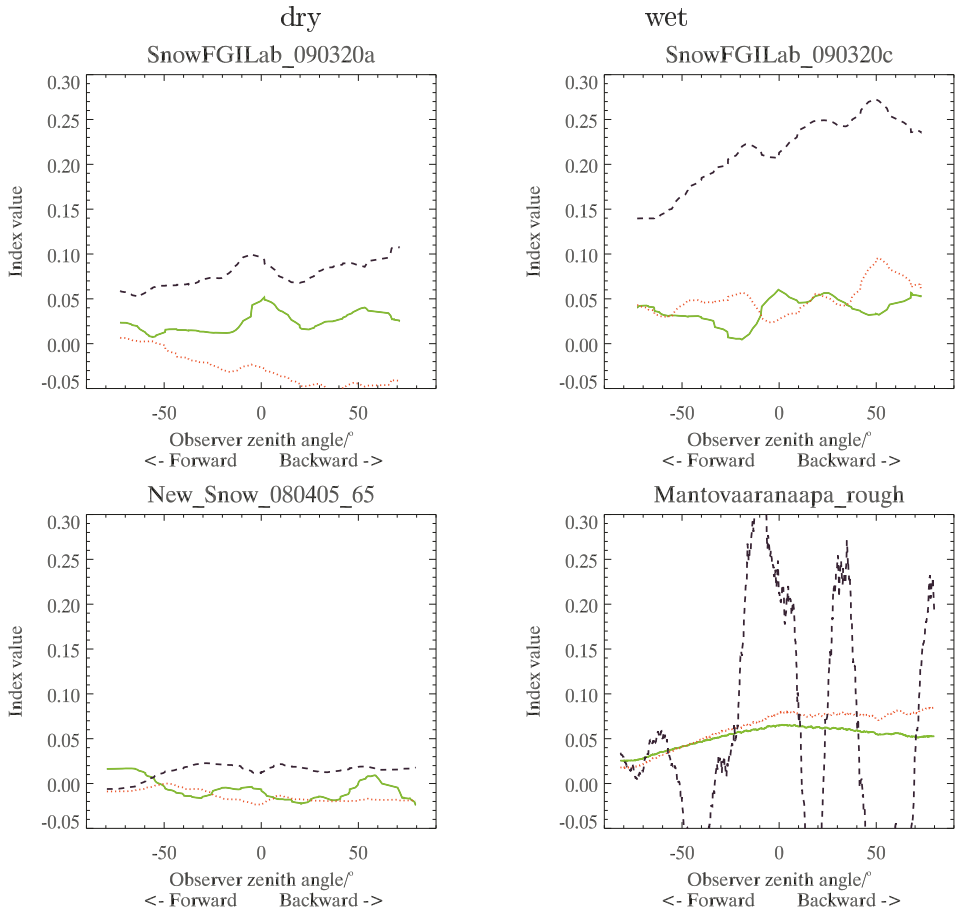
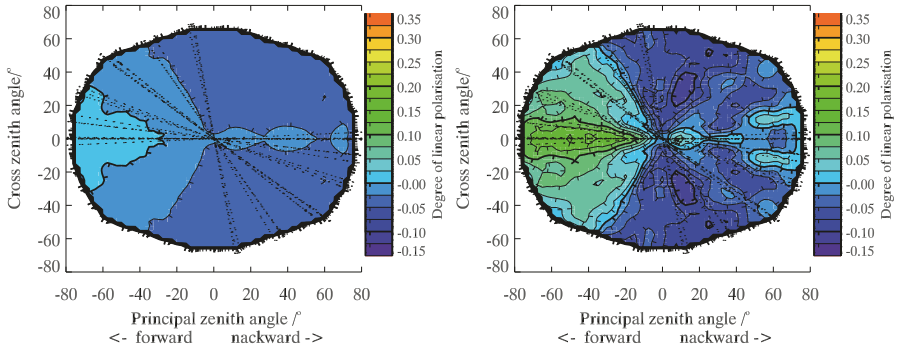
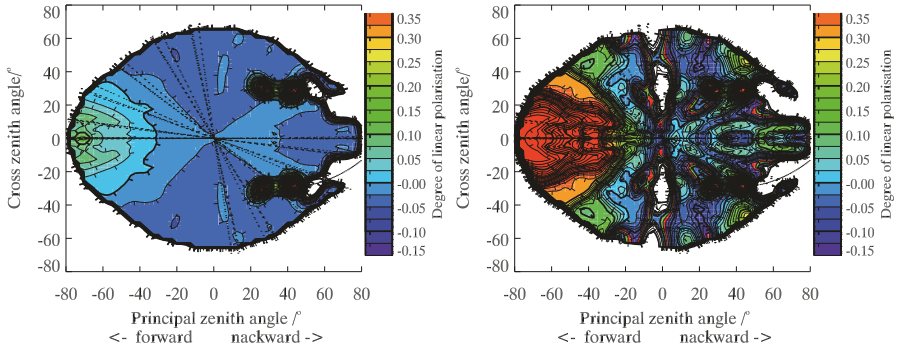


Fig. 9.26. The three wetness indices [1:green, 2:red, 3:black] as a function of the zenith angle when observing four snow samples. Dryer snows are on the left, wetter ones on the right. There is some directional dependence, and most optimum observation direction is from nadir to backwards. In sunlight measurements (Mantovaaranaapa) the index 3 contains mostly only noise, because it overlaps with atmospheric vapor absorption.

New



Dry old



Wet

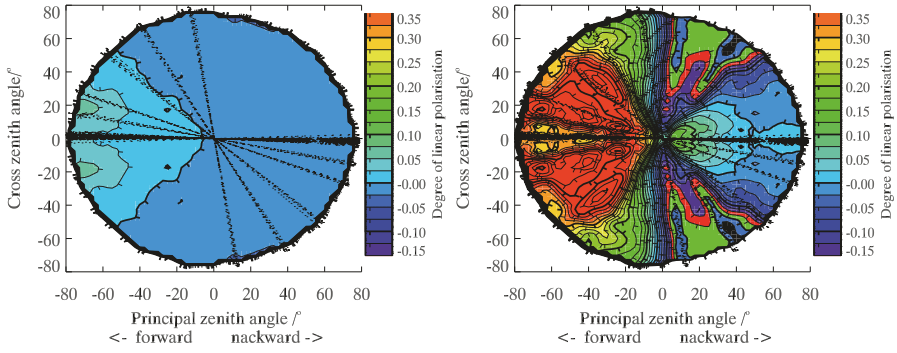


Fig. 9.27. The polarization of three snow types (from top to bottom) new snow, old snow and melting snow. On the left, a contour diagram of the linear polarization ($-Q/I$ only), in 670 nm, on the right in 1555 nm. The angle of incidence was about 65°.

Fig. 9.29. Experimenting with BRF measurements simultaneously using FIGIFIGO (on the right of the background) and Microdrone md4-200 UAV with Ricoh GR II digital camera. The pilot stands on the left of the foreground, and in the middle of the background is another spectral measurement going on.

Index