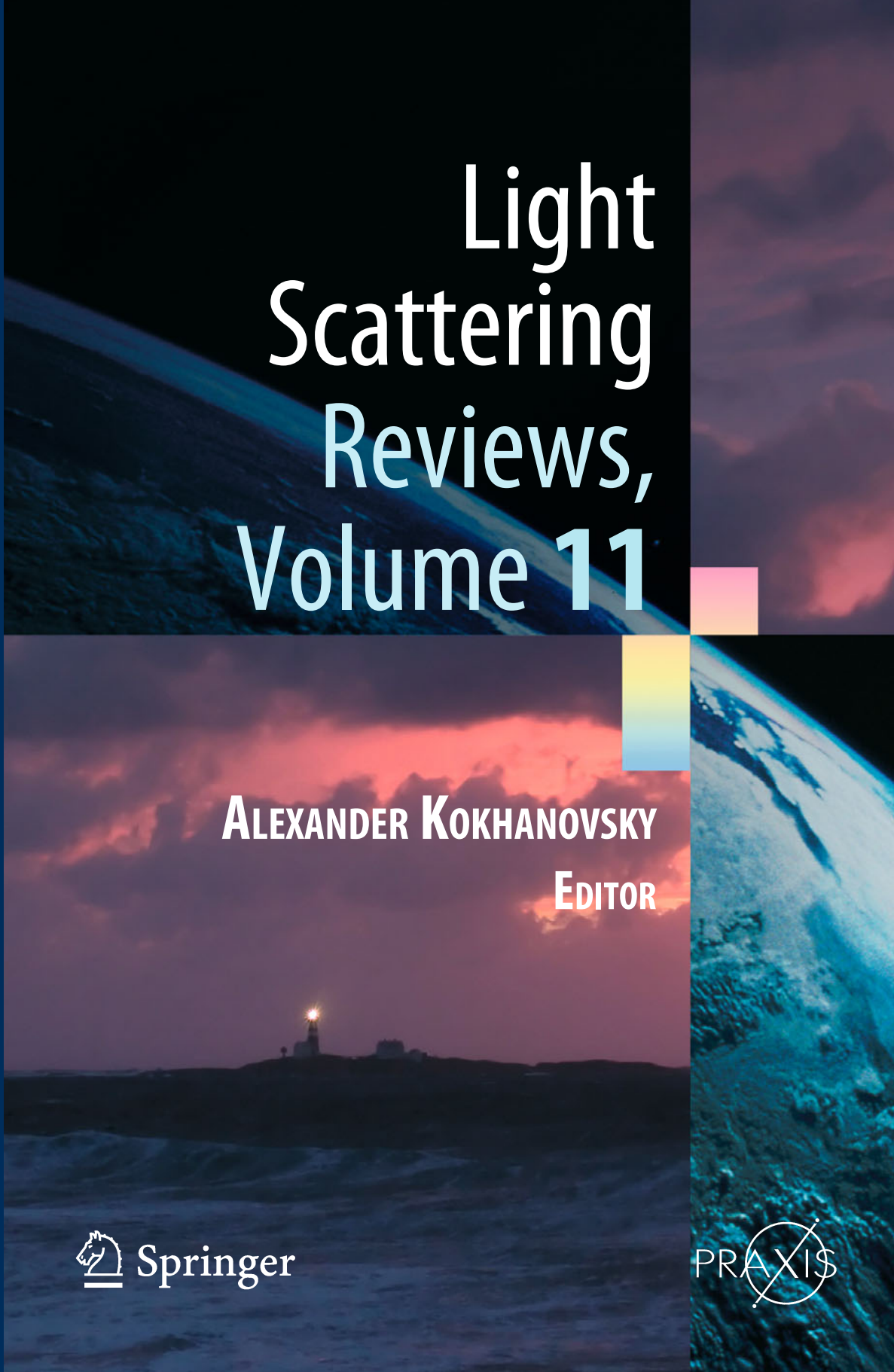

Light Scattering Reviews, Volume 11
Alexander Kokhanovsky
Editor
Springer
PRAXIS

Springer Praxis Books

More information about this series at http://www.springer.com/series/4097

Alexander Kokhanovsky
Editor

Light Scattering Reviews, Volume 11

Light Scattering and Radiative Transfer

Springer

PRAXIS

Editor
Alexander Kokhanovsky
EUMETSAT
Darmstadt
Germany

and

Moscow Engineering Physics Institute
(MEPhI)
National Research Nuclear University
Moscow
Russia

Published in association with Praxis Publishing, Chichester, UK

Springer Praxis Books
ISBN 978-3-662-49536-0 ISBN 978-3-662-49538-4 (eBook)
DOI 10.1007/978-3-662-49538-4

Library of Congress Control Number: 2016933799

© Springer-Verlag Berlin Heidelberg 2016
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer-Verlag GmbH Berlin Heidelberg

Preface

This volume of *Light Scattering Reviews* is aimed at the presentation of recent advances in radiative transfer, light scattering, and polarimetry and consists of nine chapters prepared by leading experts in respective research areas. A state-of-the-art discrete-ordinate algorithm for the transfer of monochromatic unpolarized radiation in non-isothermal, vertically inhomogeneous media, as implemented in the computer code discrete-ordinate-method radiative transfer, DISORT, is reviewed by Laszlo et al. in chapter "The Discrete Ordinate Algorithm, DISORT for Radiative Transfer". Both the theoretical background and its algorithmic implementation are covered in detail. These include features common to solutions of many radiative transfer methods, including the discrete-ordinate method, and those specific to DISORT. The common features include expansions of the phase function and the intensity into a series of Legendre polynomials and Fourier series, respectively, which transform the radiative transfer equation into a set of equations that depend only on the optical depth and the zenith angle, and the transformation of the integro-differential equations into a set of ordinary differential equations by approximating the integral in the source function by a quadrature sum. The features specific to DISORT include the reduction of the order of the standard algebraic eigenvalue problem to increase efficiency in both homogenous and particular solutions of the system of coupled ordinary differential equations, application of the scaling transformation to make the solution unconditionally stable for arbitrary large values of optical depth, application of the δ-M method to handle highly anisotropic scattering, the correction of intensities by the Nakajima–Tanaka method, and the implementation of a realistic bidirectional bottom boundary condition as realized in version 3 of DISORT. Numerical considerations that make the implementation robust and efficient are also discussed. Examples of setting up DISORT runs are shown by using test cases with increasing complexity. Brief summaries of the versions released to date are provided as well. Chapter "Community Radiative Transfer Model for Air Quality Studies" prepared by Liu and Lu presents the latest community radiative transfer model (CRTM), which is applicable for passive optical, microwave, and infrared sensors. The CRTM has

been used in operational radiance assimilations in supporting of weather forecasting and in the generation of satellite products. In this chapter, CRTM applications to assimilate aerosol optical depths derived from satellite measurements are discussed. In particular, the assimilation improves the analysis of aerosol mass concentrations. A retrieval algorithm and a retrieval product of carbon monoxide by using satellite measurements are introduced. Wei and Xu has presented the analytical solution of the time-dependent scalar and vector RTE in an infinite uniform medium with an arbitrary light scattering phase function using cumulant expansion in chapter "Analytical Solution of Radiative Transfer Using Cumulant Expansion". Analytical expressions for the exact distribution in angle and the spatial cumulants at any angle, exact up to an arbitrary high order, *n,* of photons are derived. By a cutoff at the second cumulant order, a Gaussian analytical approximate expressions of the scalar and vector photon spatial distribution is obtained as a function of the direction of light propagation and time, whose center position and half-width are always exact at arbitrary time. The center of this distribution advances and the half-width grows in time, depicting the evolution of the particle migration from near ballistic, through snake-like, and into the final diffusive regime. Contrary to what occurs in other approximation techniques, truncation of the cumulant expansion at order n is exact at that order and cumulants up to and including order n remain unchanged when contributions from higher orders are added. Various strategies to incorporate the boundary conditions in the cumulant solution are presented. The performance of the cumulant solution in an infinite and a semi-infinite medium is verified by exact numerical solutions with Monte Carlo simulations. At the end, the particular applications of the cumulant solution to RTE in biophotonics for optical imaging and in remote sensing for cloud ranging are discussed. Kolesov and Korpacheva have reviewed the radiative transfer theory in turbid media of different shapes in chapter "Radiative Transfer in Spherically and Cylindrically Symmetric Media". In particular, the authors have presented the research of radiative transfer in spherically and cylindrically symmetric media with anisotropic scattering of light. The problems of radiation transfer in an infinite homogeneous absorbing and anisotropically scattering media illuminated by a planar or point sources are considered. The relationship between the characteristics of the radiation fields in these two problems is obtained. Also an overview of the problems of radiation transfer in an infinite medium with arbitrary spherically symmetric distribution of sources is presented. The authors also discuss the structure of the radiation field in a sphere of a finite optical thickness and a spherical shell. The asymptotic expressions in the theory of radiation transfer in atmospheres with spherical symmetry are presented as well. The authors discuss the applications of the methods developed in the theory of radiative transfer in spherically symmetric media to the case of media with a cylindrical symmetry. They provide an overview of studies on the nonstationary radiative transfer in plane-parallel, spherical, and cylindrical media. Lock and Laven describe the Debye series for scattering by a sphere, a coated sphere, a multi-layer sphere, a tilted cylinder, and a prolate spheroid in chapter "The Debye Series and Its Use in Time-Domain Scattering". In electromagnetic scattering of an incident beam by a single particle possessing a reasonably high degree of

symmetry, the Debye series decomposes the partial wave scattering and interior amplitudes into the contributions of a number of intuitive physical processes. The authors comment on the meaning of the various Debye series terms, and briefly recount the method by which the formulas of ray scattering can be derived from them. They also consider time-domain scattering of a short pulse by a spherical particle and describe the way in which the time-domain scattering signature naturally separates out the various Debye series terms. Lastly, the authors show how time-domain scattering further separates a number of cooperating sub-processes present in the individual Debye series terms. Kahnert et al. discuss the models of inhomogeneous particles used in light scattering computations in chapter "Morphological Models Forinhomogeneous Particles: Light Scattering by Aerosols, Cometary Dust, and Livingcells". Light scattering by chemically heterogeneous particles with inhomogeneous internal structure is an important field of study in such diverse disciplines as atmospheric science, astronomy, and biomedical optics. Accordingly, there is a large variety of particle morphologies, chemical compositions, and dielectric contrasts that have been considered in computational light scattering studies. Depending on the intended applications, physical particle properties, and computational constraints, one can find inhomogeneous particle models ranging from simple core-shell geometries to realistic quasi-replicas of natural particles. The authors review various approaches for representing the geometry of encapsulated light-absorbing carbon aerosols, mineral dust, volcanic ash, cometary dust, and biological particles. The effects of particle inhomogeneity on radiometric properties are discussed. The authors also consider effective medium approximations, i.e., approaches that aim at avoiding the computational difficulties related to particle inhomogeneity altogether by representing such particles by a homogeneous material with an effective refractive index. Chapter "Some Wave-Theoretic Problems in Radially in Homogeneous Media" prepared by Noontaplook et al. is aimed at consideration of wave-theoretic problems in radially inhomogeneous media. The wave-theoretic aspects are based on the solution of Maxwell's equations for scattering of plane electromagnetic waves from a dielectric (or "transparent") sphere in terms of the related Helmholtz equation. There is a connection with the time-independent Schrödinger equation in the following sense: the time-independent Schrödinger equation is identical in form to the wave equation for the scalar radiation potential for TE-polarized electromagnetic waves. In regions where the refractive index is constant, it is also identical to the scalar radiation potential for TM-polarized electromagnetic waves, but with different boundary conditions than for the TE case. The authors examine scattering of the TE mode from a piecewise-uniform radial inhomogeneity embedded in an external medium (as opposed to an off-axis inclusion). The corresponding theory for the TM mode is also developed, and the well-known connection with morphology-dependent resonances (MDRs) in these contexts is noted. Kimura et al. focus on numerical approaches to deducing the light scattering and thermal emission properties of primitive dust particles in planetary systems from astronomical observations in chapter "Light Scattering and Thermal Emission by Primitive Dust Particles in Planetary systems". The particles are agglomerates of small grains with sizes

comparable to visible wavelength and compositions being mainly magnesium-rich silicates, iron-bearing metals, and organic refractory materials in pristine phases. These unique characteristics of primitive dust particles reflect their formation and evolution around main-sequence stars of essentially solar composition. The development of light scattering theories has been offering powerful tools to make a thorough investigation of light scattering and thermal emission by primitive dust agglomerates in such a circumstellar environment. In particular, the discrete dipole approximation, the T-matrix method, and effective medium approximations are the most popular techniques for practical use in astronomy. Numerical simulations of light scattering and thermal emission by dust agglomerates of submicrometer-sized constituent grains have a great potential to provide new state-of-the-art knowledge of primitive dust particles in planetary systems. What is essential to this end is to combine the simulations with comprehensive collections of relevant results from not only astronomical observations, but also in situ data analyses, laboratory sample analyses, laboratory analogue experiments, and theoretical studies on the origin and evolution of the particles. The concluding chapter "Polarimetry of Man-Made Objects" prepared by S. Savenkov is aimed at applications of environmental polarimetry. Polarimetry has already been an active area of research for about fifty years. A primary motivation for research in scatter polarimetry is to understand the interaction of polarized radiation with natural scenes and to search for useful discriminants to classify targets at a distance. In order to study the polarization response of various targets, the matrix models (i.e., 2 x 2 coherent Jones and Sinclair and 4 x 4 average power density Mueller (Stokes) and Kennaugh matrices etc.) and coherent and incoherent target decomposition techniques has been used. This comes to be the standard tools for targets characterization. Polarimetric decomposition methods allow a physical interpretation of the different scattering mechanisms inside a resolution cell. Thanks to such decompositions, it is possible to extract information related to the intrinsic physical and geometrical properties of the studied targets. This type of information is inestimable if intensity is measured only. The goal of this chapter is to explain the basics of polarimetric theory, outline its current state of the art, and review some of important applications to study the scattering behavior of various man-made and urban targets like buildings (tall and short), ships, oil rigs and spills, mines, bridges, etc. The author considers both optical range and radar polarimetry.

Darmstadt, Germany Alexander Kokhanovsky
October 2015

Contents

Part III Polarimetry

Editor and Contributors

About the Editor

Alexander Kokhanovsky graduated from the Department of Physics at the Belarusian State University, Minsk, Belarus, in 1983. He has received his Ph.D. in Optical Sciences from the B. I. Stepanov Institute of Physics, National Academy of Sciences of Belarus, Minsk, Belarus, in 1991. The Ph.D. work was on the study on light scattering properties of aerosol media and foams. His habilitation work (Main Geophysical Observatory, St. Petersburg, Russia, 2011) was devoted to the development of advanced cloud and snow remote sensing techniques based on spaceborne observations. Dr. Kokhanovsky has worked at the Atmospheric Optics group in the Laboratory of Light Scattering Media Optics at Institute of Physics in Minsk (1983–2004) and he has belonged to the SCIAMACHY/ENVISAT algorithm development team (Institute of Environmental Physics, University of Bremen, 2001–2013). He is currently working on aerosol and cloud remote sensing algorithms for the imaging polarimeter (3MI) to be launched in 2021 on board EUMETSAT Polar System-Second Generation (EPS-SG). His research interests include studies on light propagation and scattering in terrestrial atmosphere. Dr. Kokhanovsky is the author of books *Light Scattering Media Optics: Problems and Solutions* (Chichester: Springer-Praxis, 1999, 2001, 2004), *Polarization Optics of Random Media* (Berlin: Springer-Praxis, 2003), Cloud Optics (Amsterdam: Elsevier, 2006), and *Aerosol Optics* (Berlin: Springer-Praxis, 2008). He has published more than 300 papers in the field of environmental optics, radiative transfer, and light scattering.

About the Contributors

John Adam is Professor of Mathematics at the Department of Computational and Applied Mathematics (Old Dominion University, USA) and have had varied research interests, from astrophysical MHD and the related singular differential equations, to modeling tumor growth and wound healing. His current area of research is in theoretical meteorological optics, specifically ray and wave propagation in radially inhomogeneous media, with potential applications to the study of certain types of metamaterials. He has written several books published by Princeton University Press (see http://www.odu.edu/∼jadam for details), including *Mathematics in Nature: Modeling Patterns in the Natural World*.

Wei Cai is a Senior Scientist at the Institute of the Ultrafast Laser and Spectroscopies, City College of City University of New York. He received his Bachelor degree in Physics at Quinghua University, Beijing, China and later a Ph.D. degree in Physics at University of Houston in 1985. He has published a total of 63 papers in Physical Review Letters, Physical Review, and other peer-referred journals. Dr. Cai hold six U.S. patents and is the co-author of the book titled "Random Processes in Physics and Finance" (2006, Oxford Press).

Michael Kahnert is a Senior Scientist at the Research Department of the Swedish Meteorological and Hydrological Institute, and an Adjunct Professor at the Department of Earth and Space Sciences of Chalmers University of Technology in Gothenburg. His research interests cover fundamental questions and numerical methods in electromagnetic scattering theory, applications to aerosols, ice clouds, and cosmic dust, radiative transfer in planetary atmospheres, as well as inverse modeling, remote sensing, chemical data assimilation, air pollution modeling, and aerosol-cloud coupling. He

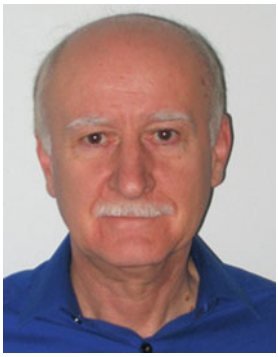
received his Ph.D. in 1998 at the University of Alaska Fairbanks (USA), and his Habilitation/senior Lecturer Expertise in 2010 at Chalmers University of Technology (Sweden).

Hiroshi Kimura is currently a Research Fellow at the Graduate School of Science of Kobe University in Japan. His research interest lies in the life cycle of cosmic dust in order to answer the question as to whether cosmic dust is the origin of life in the universe. He has conducted research on physical and chemical processes of dust occurring in interplanetary space, cometary comae and tails, the surfaces of airless bodies, debris disks, protoplanetary disks, and the interstellar medium. He has been leading the Scientific Organization Committee (SOC) and the Local Organizing Committee (LOC) of a series of the international meeting on Cosmic Dust, since he established the SOC and the LOC in 2006.

Alexander Kolesov is a Professor of the Department of General Mathematics and Informatics of the St. Petersburg State University. Kolesov is one of the representatives of the scientific school of theoretical astrophysics, established in St. Petersburg (Leningrad) University, by the academician Ambartsumian and headed by academician Sobolev. His research interests relate to the theory of multiple scattering of light in anisotropic scattering media with flat, cylindrical, or spherical symmetry. In 1988 he was awarded the degree of Doctor of Physical and Mathematical Sciences for the development of the theory of radiation transport in a spherically symmetric media. In recent years he studies the nonstationary radiation fields generated by a momentary point source.

Ludmilla Kolokolova is a Senior Research Scientist at the Astronomy Department of the University of Maryland, College Park. Her main scientific interest is remote sensing of dust in Solar system and beyond using photopolarimetric techniques. Her research is focused on laboratory and computer modeling of light scattering by complex particles and regolith surfaces and interpretation of ground-based and space mission observations. She is also the manager of the NASA Planetary Data System Small Bodies Node and is responsible for long-term preservation and archiving of the data collected by space missions to small bodies (comets, asteroids, TNOs, satellites of planets).

Natalia Kropacheva is an Associate Professor of the Department of General Mathematics and Informatics of the St. Petersburg State University, Candidate of Physical and Mathematical Sciences, specialist in the field of modeling complex systems. In recent years the range of her scientific interests included the study of nonstationary processes of radiation transfer in spherical and cylindrical symmetric media, illuminated by a momentary point source.



Istvan Laszlo is a Physical Scientist at the Center for Satellite Applications and Research of the National Oceanic and Atmospheric Administration, and an Adjunct Professor at the Department of Atmospheric and Oceanic Science of the University of Maryland in College Park, MD, USA. He was an elected member of the International Radiation Commission. He has published over 50 peer-reviewed scientific articles, and two book chapters. His research interests include radiative transfer, radiation budget, and remote sensing of aerosols. Laszlo earned an M.Sc. in Astronomy and a Ph.D. in Earth Sciences from the Eötvös Loránd University, Budapest, Hungary.

Philip Laven is currently the Chairman of DVB, a not-for-profit organization that develops technical standards for digital TV (see http://www.dvb.org). His professional life has been in broadcast engineering, first with the BBC and then in the European Broadcasting Union. In his spare time, he investigates meteorological optics and other light scattering phenomena.

Jérémy Lebreton is a Postdoctoral Researcher at the California Institute of Technology in Pasadena, California. He obtained his Ph.D. from the Institute of Planetology and Astrophysics of Grenoble, France, where he acquired expertise in observations and modeling of planetary systems around nearby stars. He works on debris disks and planets at the NASA Exoplanet Science Institute using detailed radiative transfer models and simulations based on far-infrared imaging and photometry, optical and mid-infrared interferometry, and coronagraphic imaging.

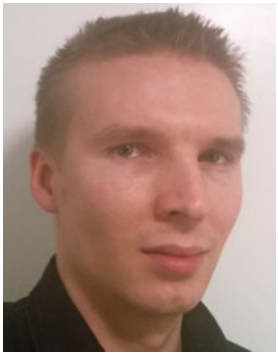
Aigen Li is a Professor of Astrophysics at the University of Missouri in the USA. Since he started his graduate study in 1994 under the supervision of J. Mayo Greenberg at Leiden University in The Netherlands, he has been studying the physics and chemistry of the interstellar medium and interstellar dust, comets, planet-forming dust disks, dust-making evolved stars, and dust in external galaxies, active galactic nuclei, and gamma-ray bursts.

Quanhua Liu is a Physical Scientist at the NOAA/NESDIS and is leading the soundings team at NOAA/STAR. The sounding system utilized satellite-based microwave observations and infrared hyper-spectral measurements to acquire vertical profiles of atmospheric temperature, water vapor, ozone, CO, CH_4, CO_2, and others chemical species. Quanhua Liu studied the infrared hyper-spectral sensor and the community radiative transfer model (CRTM). The CRTM has been operationally supporting satellite radiance assimilation for weather forecasting. The CRTM also supports JPSS/NPP and GOES-R missions for instrument calibration, validation, long-term trend monitoring, and satellite retrieved products. He received a B.S. from Nanjing University of Information Science and Technology (former Nanjing Institute of Meteorology), China, a Master's Degree in Physics from the Chinese Academy of Science, and a Ph.D. in Meteorology and Remote Sensing from the University of Kiel, Germany.

James Lock received his Ph.D. in Theoretical Physics in 1974 from Case Western Reserve University in Cleveland OH, USA. Following a 40-year teaching career first in the physics department of CWRU and then in the Physics Department of Cleveland State University, he is currently Professor Emeritus at CSU. His main research interest is single-scattering of a plane wave or a shaped beam from a particle having a relatively high degree of symmetry. He is especially interested in the wide variety of beautiful phenomena that dominate scattering in the short-wavelength limit.

Cheng-Hsuan Lu is a Research Associate at Atmospheric Sciences Research Center (ASRC), University at Albany, State University of New York. Dr. Lu's current research focuses on quantifying the distributions of tropospheric aerosols, its impact on weather forecasts and climate predictions, and improving aerosol forecasts through the assimilation of satellite and in situ aerosol observations. She led the efforts to develop the global aerosol forecast system at NOAA/NCEP before joining ASRC in September 2014 and remains actively involved with the global aerosol prediction community. She received a B.S. from

National Central University, Taiwan, and a Master's Degree and a Ph.D. in Atmospheric Sciences from University at Albany, State University of New York.



Johannes Markkanen received his D.Sc. (Tech.) degree in electromagnetics from the Aalto University, Finland, in 2013. Currently he is a Postdoctoral Researcher at the Department of Physics, University of Helsinki, Finland. He works with the planetary science research group and his research interests include computational methods in electromagnetics (particularly integral equation methods, regularization techniques, and fast algorithms), multiple-scattering theory, and electromagnetic modeling of natural and artificial materials.

Timo Nousiainen has studied the single-scattering modeling of various types of nonspherical, atmospheric particles for over 20 years, during which he has authored or co-authored over 60 peer-reviewed original research papers and three reviews. He has contributed particularly to research on mineral dust particles, with secondary emphasis on tropospheric ice crystals. For the last 2 years, he has mainly written proposals on various topics and managed weather-radar-related service development. He is currently working in the Radar and Space Technology group at the Finnish Meteorological Institute.

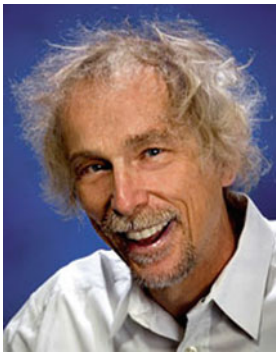
Umaporn Nuntaplook received her Ph.D. in Computational and Applied Mathematics from Old Dominion University, USA, in 2013. Currently she is a Lecturer at the Department of Mathematics, Mahidol University, Bangkok, Thailand. Her research interests are applying the computational methods to the problems in the fields of mathematical physics including the topics in electromagnetic, acoustic, and potential scattering theory.

Michael A. Pohrivchak is working at Naval Research Laboratory (Washington, D.C., USA). His primary focus was on electromagnetic theory with an emphasis on nonhomogeneous media. Currently, he is working on the propagation of electromagnetic waves through the atmosphere. This requires obtaining an accurate solution to the parabolic wave equation which allows for the prediction of the path loss between two points.

Sergey N. Savenkov received his Ph.D. in Radiophysics 1996 and his Doctor of Science degree in Optics and Laser Physics in 2013 from Taras Shevchenko National University of Kyiv. His research interests are focused on the use of polarization of electromagnetic radiation in remote sensing, crystal optics, biomedical optics, and laser/radar technology. He is currently employed at Faculty of Radio Physics, Electronics and Computer Systems of Taras Shevchenko National University of Kyiv. Specific topics are the methods of structural and anisotropy property analysis of objects contained in their scattering matrices, development of the Mueller matrix measurement methods, optimization polarimetric sensors for remote sensing applications, and use of polarimetry to improve imaging in turbid media.

Knut Stamnes is a Professor in the Department of Physics and Engineering Physics, and Director of the Light and Life Laboratory at Stevens Institute of Technology, Hoboken New Jersey. His research interests include radiative transfer, remote sensing, and climate-related studies. He has published more than 200 research papers and two textbooks (Radiative Transfer in the Atmosphere and Ocean, Cambridge, 1999; Radiative Transfer in Coupled Environmental Systems, Wiley-VCH, 2015). He is a fellow of OSA, a member of SPIE, and was elected member of the Norwegian Academy of Technological Sciences in 2009. Stamnes earned his B.Sc. and M.Sc. in Physics from the University of Oslo, Norway, and his Ph.D. from the University of Colorado, USA.

Si-Chee Tsay is a Senior Physical Scientist in NASA Climate and Radiation Laboratory and serves as EOS/Terra deputy project scientist since 2000. Besides theoretical radiative transfer, his research experience includes science planning and operating remote sensing instruments and in situ probes from suborbital and ground-based platforms in field experiments to study aerosol and cloud properties, and to aid in the development of retrieval algorithms. He has published over a hundred of research papers and awarded *NASA Exceptional Scientific Achievement Medal* in 2002 and *NASA Outstanding Leadership Medal* in 2008, among numerous achievement awards. Tsay earned his B.S. in Atmospheric Sciences from the National Taiwan University, Taiwan, and M.S. and Ph.D. in Atmospheric Sciences from the University of Alaska.



Warren Wiscombe graduated from MIT with a B.S. in Physics, then from CalTech with a Ph.D. in Applied Mathematics. He joined the ARPA Climate Dynamics Program in 1971 and helped found the new field of climate science. He specialized in atmospheric radiation with particular application to clouds. He worked at NCAR in the 1970s, then at NYU as a Professor, and then at NASA Goddard for 30 years before retiring in 2013. His work until 1990 was largely theoretical, but increasingly afterward became measurement-oriented, leading ultimately to his capstone job as DOE ARM Chief Scientist.

Min Xu obtained his Ph.D. in Physics from the City University of New York in 2001 and is currently Associate Professor of Physics at Fairfield University, CT, USA. He has been actively involved in biophotonics since 2001 with over 50 peer-reviewed journal papers, books and book chapters, and five patents. His research focuses on random photonics, and spectroscopy and tomography in biomedical applications. In particular, he is interested in mathematical modeling and applications of multiply scattered light in complex systems. Dr. Xu is currently developing novel mesoscopic and microscopic quantitative imaging techniques for lighting up the static structure and dynamic

processes in biological systems; and photonic cancer screening, diagnosis, and prognosis. He pioneered the unified Mie and fractal model for light scattering by complex systems such as cell and tissue, and the electric field Monte Carlo method to simulate coherence phenomenon of multiple scattered light. He is a Senior Member of OSA.

Contributors

John A. Adam Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA, USA

Wei Cai Physics Department, The City College of New York, New York, NY, USA

Michael Kahnert Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

Hiroshi Kimura Graduate School of Science, Kobe University, C/O CPS (Center for Planetary Science), Kobe, Japan

Alexander Kolesov Department of General Mathematics and Informatics, St. Petersburg State University, St. Petersburg, Russia

Ludmilla Kolokolova Planetary Data System Group, Department of Astronomy, University of Maryland, College Park, MD, USA

Nataliya Kropacheva Department of General Mathematics and Informatics, St. Petersburg State University, St. Petersburg, Russia

I. Laszlo Center for Satellite Applications and Research, National Oceanic and Atmospheric Administration, College Park, MD, USA

Philip Laven 9 Russells Crescent, Horley, UK

Jérémy Lebreton Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA, USA; NASA Exoplanet Science Institute, California Institute of Technology, Pasadena, CA, USA

Aigen Li Department of Physics and Astronomy, University of Missouri, Columbia, MO, USA

Quanhua Liu Center for Satellite Applications and Research, National Oceanic and Atmospheric Administration, College Park, MD, USA

James A. Lock Department of Physics, Cleveland State University, Cleveland, OH, USA

Cheng-Hsuan Lu Atmospheric Sciences Research Center, State University of New York, Albany, NY, USA

Johannes Markkanen Department of Physics, University of Helsinki, Helsinki, Finland

Timo Nousiainen Finnish Meteorological Institute, Helsinki, Finland

Umaporn Nuntaplook Faculty of Science, Department of Mathematics, Mahidol University, Bangkok, Thailand

Michael A. Pohrivchak Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA, USA; Naval Research Laboratory, Washington, DC, USA

Sergey N. Savenkov Faculty of Radio Physics, Electronics and Computer Systems, Taras Shevchenko National University of Kyiv, Kiev, Ukraine

K. Stamnes Department of Physics and Engineering Physics, Stevens Institute of Technology, Hoboken, NJ, USA

S.-C. Tsay Climate and Radiation Laboratory, Code 613, NASA Goddard Space Flight Center, Greenbelt, MD, USA

W.J. Wiscombe Climate and Radiation Laboratory, Code 613, NASA Goddard Space Flight Center, Greenbelt, MD, USA

Min Xu Physics Department, Fairfield University, Fairfield, CT, USA

Part I
Radiative Transfer

The Discrete Ordinate Algorithm, DISORT for Radiative Transfer

Istvan Laszlo, Knut Stamnes, Warren J. Wiscombe and Si-Chee Tsay

Abstract The discrete ordinate method for the transfer of monochromatic unpolarized radiation in non-isothermal, vertically inhomogeneous media, as implemented in the computer code Discrete-Ordinate-Method Radiative Transfer, DISORT, is reviewed. Both the theoretical background and its algorithmic implementation are covered. Among others, described are the reduction of the order of the standard algebraic eigenvalue problem to increase efficiency in both the homogenous and particular solutions of the system of coupled ordinary differential equations, application of the scaling transformation to make the solution unconditionally stable for arbitrary large values of optical depth, application of the δ-M method to handle highly anisotropic scattering, the correction of intensities by the Nakajima-Tanaka method, and the implementation of a realistic bidirectional bottom boundary. Numerical considerations that make the implementation robust and efficient are also discussed. Examples of setting up DISORT runs are shown by using test cases with increasing complexity. Brief summaries of the versions released to date are provided, as well.

I. Laszlo (✉)
Center for Satellite Applications and Research, National Oceanic and Atmospheric Administration, 5825 University Research Court, College Park, MD 20740, USA
e-mail: Istvan.Laszlo@noaa.gov

K. Stamnes
Department of Physics and Engineering Physics, Stevens Institute of Technology, Hoboken, NJ 07030, USA
e-mail: Knut.Stamnes@stevens.edu

W.J. Wiscombe · S.-C. Tsay
Climate and Radiation Laboratory, Code 613, NASA Goddard Space Flight Center, 8800 Greenbelt Road, Greenbelt, MD 20771, USA
e-mail: warren.j.wiscombe@gmail.com

S.-C. Tsay
e-mail: si-chee.tsay@nasa.gov

© Springer-Verlag Berlin Heidelberg 2016

A. Kokhanovsky (ed.), *Light Scattering Reviews, Volume 11*,
Springer Praxis Books, DOI 10.1007/978-3-662-49538-4_1

1 Introduction

Studies of propagation of electromagnetic radiation in a medium (e.g., stellar and planetary atmospheres) require the solution of an equation, the radiative transfer equation, which mathematically describes the interaction between emission, absorption, and scattering by which the medium affects the transfer of radiation. One such solution is the discrete ordinate approximation; a systematic development of which is presented by Chandrasekhar (1960). The strength of the approximation is in the transformation of an integro-differential equation describing radiative transfer to a system of ordinary differential equations for which solutions in terms of eigenvectors and eigenvalues can be found. Computer implementation of the discrete ordinate solutions proposed by investigators (e.g., Chandrasekhar 1960; Liou 1973; Asano 1975), however, had numerical difficulties as discussed by Stamnes and Swanson (1981), who also showed a way to overcome these difficulties. The discrete ordinate method has gained considerable popularity after the publication of the paper by Stamnes et al. (1988a) that presented a detailed summary of treating numerical ill-conditioning, computation of the eigenvalues and eigenvectors, efficient inversion of the matrix needed for determining the constants of integration, and especially after its implementation in the computer code Discrete Ordinate Method Radiative Transfer, or DISORT, in 1988 was made readily available to the public.

DISORT is a discrete ordinate algorithm for monochromatic unpolarized radiative transfer in non-isothermal, vertically inhomogeneous, but horizontally homogeneous media. It can treat thermal emission, absorption, and scattering with an arbitrary phase function covering the electromagnetic spectrum from the ultraviolet to radio. The medium may be driven by parallel or isotropic diffuse radiation incident at the top boundary, by internal thermal sources and thermal emission from the boundaries, as well as by bidirectional reflection at the surface. It calculates intensities (radiances), fluxes (irradiances), and mean intensities at user-specified angles and levels.

Our goal is to review the discrete ordinate approximation as it is implemented in DISORT. The primary source used in this review is the DISORT Report v1.1 (Stamnes et al. 2000). Most of the material is taken from that report with little or no modification. However, some parts, e.g., the treatment of the bidirectionally reflecting lower boundary, are expanded on. We also include recent advances that appeared in Lin et al. (2015), and which were not present in the v1.1 Report. We first describe the theoretical basis for DISORT, and then discuss numerical considerations that must be dealt with in order to make the implementation robust and efficient. Next, taking from the many test cases provided with the code, we show examples of how to correctly set up a DISORT run, and finally, we provide brief summaries of the versions released to date.

2 Equation of Transfer in DISORT

We consider a plane-parallel horizontally homogeneous medium (a slab), the optical properties of which are characterized by its optical thickness (defined as the difference between the optical depth at the bottom and that at the top), the single-scattering albedo ω, and the scattering phase function P. As defined, ω gives the fraction of an incident beam which is scattered by an infinitesimal volume inside the medium, while P describes how much of the radiation incident from a given direction is scattered by that volume into another direction, that is, the angular scattering pattern.

Location in the medium is specified by one vertical coordinate, measured in optical depth units (τ) from the top down. Directions are described by two angular coordinates, zenith and azimuth angles (Fig. 1). Polar (zenith) angles (θ) are measured from the upward direction: straight up is 0° and straight down is 180°. In the rest of the discussion we use the cosine of the polar angle (μ) instead of θ to denote the polar direction. According to the above convention all upward directions have positive polar angle cosines ($+\mu$), while downward-directed intensities have negative polar angle cosines ($-\mu$). The exception is the cosine of the incident beam angle (μ_0) which, for historical reasons, is taken positive. Azimuth angles (ϕ) are measured in an absolute frame of reference between 0° and 360°. They can be measured either clockwise or counterclockwise from the zero azimuth when viewed downward from zenith. However, when a choice has been made it must be applied consistently for all directions (upward, downward, incident, reflection). We note that according to this definition the relative azimuth angle of sunglint is 0°.

The transfer of monochromatic radiation through the medium, subject to internal thermal emission in local thermodynamic equilibrium and illuminated at the top boundary by a parallel beam in the direction μ_0, ϕ_0, is described by the following pair of equations:

$$I_{\text{direct}}(\tau) = F_0 \mathrm{e}^{-\tau/\mu_0} \delta(\mu - \mu_0)\delta(\phi - \phi_0), \tag{1}$$

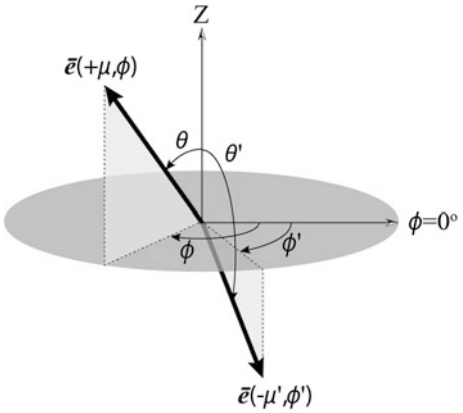


Fig. 1 Definition of upward, $\bar{e}(+\mu, \phi)$ and downward, $\bar{e}(-\mu, \phi)$ directions in DISORT. Polar (zenith) angles θ are measured from the upward normal pointing to Z. Azimuth angles ϕ are measured in a plane perpendicular to the upward normal, and in relative to the reference direction $\phi = 0°$

$$\mu \frac{\mathrm{d}I(\tau,\mu,\phi)}{\mathrm{d}\tau} = I(\tau,\mu,\phi) - S(\tau,\mu,\phi). \tag{2}$$

Here I_{direct} is the intensity of the direct beam at vertical optical depth τ, F_0 is the flux (irradiance) of the parallel beam normal to the direction of incidence at the top boundary, and δ is the delta function in units of per steradian. I is the diffuse specific intensity at τ in a cone of unit solid angle along direction μ, ϕ, and S is the "source function." S is the sum of the radiation scattered into the direction μ, ϕ from all other directions μ', ϕ', the "pseudobeam" source term $Q^{(\mathrm{beam})}$, and the internal thermal source $Q^{(\mathrm{thermal})}$ characterized by the Planck function $B(T)$ at temperature T at optical depth τ (cf. Stamnes et al. 2000):

$$\begin{aligned} S(\tau,\mu,\phi) = \frac{\omega(\tau)}{4\pi} \int_0^{2\pi} \mathrm{d}\phi' \int_{-1}^{1} P(\tau,\mu,\phi;\,\mu',\phi') I(\tau,\mu',\phi') \mathrm{d}\mu' \\ + Q^{(\mathrm{beam})}(\tau,\mu,\phi) + Q^{(\mathrm{thermal})}(\tau), \end{aligned} \tag{3}$$

where

$$\begin{aligned} Q^{(\mathrm{beam})}(\tau,\mu,\phi) &= \frac{\omega(\tau)}{4\pi} P(\tau,\mu,\phi;\,-\mu_0,\phi_0) F_0\, \mathrm{e}^{-\tau/\mu_0}, \\ Q^{(\mathrm{thermal})}(\tau) &= \{1-\omega(\tau)\} B[T(\tau)]. \end{aligned} \tag{4}$$

Apart from the polar angles defining the direction all other quantities in (1)–(4) depend on the wavelength of radiation. The wavelength dependence is assumed to be understood and is omitted from the equations.

Equation (1) gives the solution for the transfer of the direct beam radiation. It says that I_{direct} decreases exponentially with the pathlength τ/μ_0, and it is nonzero only in the direction μ_0, ϕ_0. The solution of (2) provides the diffuse radiation propagating in the direction μ, ϕ at the optical depth τ. In the remainder of the document we describe the solution of (2) using the discrete ordinate method as implemented in DISORT. The solution, as we show below, is comprised of essentially three steps: (1) transforming (2) into a set of radiative transfer equations which are functions of the vertical coordinate τ and the angular coordinate μ only (separation of azimuth dependence); (2) transforming the integro-differential equations into a system of ordinary differential equations; and (3) solving the system of ordinary differential equations using robust linear algebra solvers.

2.1 *Radiative Transfer Equation Uncoupled in Azimuth*

The scatterers within the medium are assumed to have random orientations; thus, ω does not explicitly depend on the direction of the incident beam, and P depends only on the angle between the incident and scattered beam (the scattering angle, Θ),

that is, $P(\tau, \mu, \phi; \mu', \phi') = P(\tau, \cos\Theta)$ where, from the cosine law of spherical trigonometry, $\cos\Theta = \mu\mu' + \sqrt{(1-\mu^2)(1-\mu'^2)}\cos(\phi - \phi')$. With this restriction on the form of P, we expand the phase function P into a series of Legendre polynomials P_ℓ with $2M$ terms ($\ell = 0, \ldots, 2M-1$) (Chandrasekhar 1960, Chap. 1, Eq. 33; Thomas and Stamnes 1999, Eq. 6.28).

$$P(\tau, \cos\Theta) = \sum_{\ell=0}^{2M-1} (2\ell+1)\, g_\ell(\tau) P_\ell(\cos\Theta), \tag{5}$$

where the expansion coefficients g_ℓ are given by

$$g_\ell(\tau) = \frac{1}{2}\int_{-1}^{+1} P_\ell(\cos\Theta) P(\tau, \cos\Theta)\mathrm{d}(\cos\Theta). \tag{6}$$

In DISORT we require the phase function to be normalized to unity, so $g_0 = 1$. The g's generally decrease monotonically, so we can expect that a finite number of terms $2M$ in the expansion is sufficient. However, for highly asymmetric phase functions (e.g., for clouds) the g's often decrease very slowly, and several hundred terms may be necessary in (5) to adequately represent the phase function (in Sect. 6.1 we show how DISORT mitigates this problem).

Applying the addition theorem for spherical harmonics (Chandrasekhar 1960, Chap. 6, Eq. 86; Thomas and Stamnes 1999, Eq. 6.30) to (5) we obtain

$$P(\tau, \mu, \phi; \mu', \phi') = \sum_{m=0}^{2M-1} (2 - \delta_{0m}) \left\{ \sum_{\ell=m}^{2M-1} (2\ell+1) g_\ell(\tau) \Lambda_\ell^m(\mu) \Lambda_\ell^m(\mu') \right\} \cos m(\phi - \phi'). \tag{7}$$

Here Λ_ℓ^m are the normalized associated Legendre polynomials related to the associated Legendre polynomials P_ℓ^m by

$$\Lambda_\ell^m(\mu) = \sqrt{\frac{(\ell-m)!}{(\ell+m)!}} P_\ell^m(\mu).$$

Since (7) is essentially a Fourier expansion of P in azimuth, we may similarly expand the intensity in a Fourier cosine series (Chandrasekhar 1960, Chap. 6, Eq. 91; Thomas and Stamnes 1999, Eq. 6.34):

$$I(\tau, \mu, \phi) = \sum_{m=0}^{2M-1} I^m(\tau, \mu) \cos m(\phi_0 - \phi). \tag{8}$$

Substitution of this equation, as well as (3) and (7) into the radiative transfer equation (2) splits it into $2M$ independent integro-differential equations, one for each azimuthal intensity component I^m:

$$\mu \frac{\mathrm{d}I^m(\tau,\mu)}{\mathrm{d}\tau} = I^m(\tau,\mu) - S^m(\tau,\mu), \quad (m = 0, 1, \ldots, 2M-1), \tag{9}$$

where the source function S is given by

$$S^m(\tau,\mu) = \int_{-1}^{1} D^m(\tau,\mu,\mu') I^m(\tau,\mu')\mathrm{d}\mu' + Q^m(\tau,\mu). \tag{10}$$

The symbols D^m and Q^m are defined by

$$D^m(\tau,\mu,\mu') = \frac{\omega(\tau)}{2} \sum_{\ell=m}^{2M-1} (2\ell+1) g_\ell(\tau) \Lambda_\ell^m(\mu) \Lambda_\ell^m(\mu'), \tag{11}$$

$$Q^m(\tau,\mu) = X_0^m(\tau,\mu) \mathrm{e}^{-\tau/\mu_0} + \delta_{m0}\, Q^{(\mathrm{thermal})}(\tau), \tag{12}$$

where

$$X_0^m(\tau,\mu) = \frac{\omega(\tau)\, F_0}{4\pi} (2 - \delta_{m0}) \sum_{\ell=m}^{2M-1} (-1)^{\ell+m} (2\ell+1) g_\ell(\tau) \Lambda_\ell^m(\mu) \Lambda_\ell^m(\mu_0), \tag{13}$$

and δ_{m0} is the Kronecker delta

$$\delta_{m0} = \begin{cases} 1 & \text{if } m = 0, \\ 0 & \text{otherwise.} \end{cases}$$

The above procedure transforms (2) into a set of equations (9) which do not depend on the azimuth angle (ϕ). It also uncouples the various Fourier components I^m in (9); that is, I^m does not depend on any I^{m+k} for $k \neq 0$.

Using the same number of terms ($2M$) in the Fourier expansion of intensity (8) as in the Legendre polynomial expansion of the phase function (7) is not accidental. To explain why, let us consider the case when the number of terms is different: $2M$ in the expansion of P and $2K$ in the expansion of I, and assume that $K > M$. Substitution of the series expansions into (2) results in an equation containing terms that are proportional to integrals of the type

$$\int_0^{2\pi} \sum_{m=0}^{2M-1} \sum_{k=0}^{2M-1} \cos mx \, \cos kx \, \mathrm{d}x + \int_0^{2\pi} \sum_{m=0}^{2M-1} \sum_{k=2M}^{2K-1} \cos mx \, \cos kx \, \mathrm{d}x.$$

In this expression the second term is zero because m runs from 0 to $2M - 1$ and k runs from $2M$ to $2K - 1$; that is, for each term $k \neq m$, and in this case the integral is zero. The same argument holds for the case $M > K$, in which case the remaining (nonzero) integral has both k and m running from 0 to $2K - 1$. As a consequence the

expansion of the phase function and that of the intensity will have the same number of terms; and that number is determined by the smaller of M and K. We note that the integrals with $k \neq m \neq 0$ in the first term are also zero, which leads to the uncoupling of the various Fourier components mentioned above.

We note that when the lower boundary of the medium is characterized by specular reflection, for example, at the bottom of atmosphere in the atmosphere–ocean system, specular reflection can be included by adding a reflected beam source term to (12)

$$X_0^m(\tau, \mu)\, \rho_s(-\mu_0, \phi_0)\, \mathrm{e}^{-(2\tau_L - \tau)/\mu_0},$$

where ρ_s is the specular reflection function. This term is, however, not included in the versions of DISORT reviewed here.

3 Discrete Ordinate Approximation—Matrix Formulation

The steps presented so far are common to many approaches used to solve (2). What sets the discrete ordinate method apart from these, and gives its name, is the next step. In this step, we approximate the integral in (10) by a quadrature sum. For later convenience, we choose even-order quadrature angles $2N$ in the sum so that we have the same number of polar angle cosines for $+\mu$ as for $-\mu$. Substitution of the integral with a quadrature sum transforms the integro-differential equation (9) into the following system of ordinary differential equations (cf. Stamnes and Dale 1981; Stamnes and Swanson 1981)

$$\mu_i \frac{\mathrm{d}I^m(\tau, \mu_i)}{\mathrm{d}\tau} = I^m(\tau, \mu_i) - S^m(\tau, \mu_i), \quad (i = \pm 1, \ldots, \pm N). \tag{14}$$

Each μ_i is called a "stream", and (14) represents a "$2N$—stream approximation".

Writing (10) in quadratured form, S^m becomes a linear combination of I^m values at all quadrature angles μ_j ($j = \pm 1, \ldots, \pm N$),

$$S^m(\tau, \mu_i) = \sum_{\substack{j=-N \\ j\neq 0}}^{N} w_j D^m\left(\tau, \mu_i, \mu_j\right) I^m\left(\tau, \mu_j\right) + Q^m(\tau, \mu_i). \tag{15}$$

This approach makes the system coupled in i, but not in m.

In DISORT we draw the μ_i from a Gaussian quadrature rule for [0, 1] and have them mirror symmetric ($\mu_{-i} = -\mu_i$, where $\mu_i > 0$) with weights $w_{-i} = w_i$. This scheme is the so-called "Double-Gauss" quadrature rule suggested by Sykes (1951) in which Gaussian quadrature is applied separately to the half-ranges $-1 < \mu < 0$ and $0 < \mu < 1$. The main advantage is that even-order quadrature points are distributed symmetrically around $|\mu| = 0.5$ and clustered both toward $|\mu| = 1$ and $\mu = 0$,

whereas in the Gaussian scheme for the complete range $-1 < \mu < 1$, they are clustered only toward $\mu = -1$ and $\mu = +1$. The Double-Gauss clustering toward $\mu = 0$ will give superior results anywhere the intensity varies rapidly across $\mu = 0$, especially at the boundaries where the intensity is often discontinuous at $\mu = 0$. Another advantage is that upward and downward fluxes, as we show later, are obtained immediately without any further approximation. The use of Gaussian quadrature also guarantees that the phase function is normalized, i.e.,

$$\sum_{\substack{j=-N \\ j\neq 0}}^{N} w_j D^0(\tau, \mu_i, \mu_j) = \sum_{\substack{i=-N \\ i\neq 0}}^{N} w_i D^0(\tau, \mu_i, \mu_j) = \omega(\tau), \tag{16}$$

implying that energy is conserved in the computation (cf. Wiscombe 1977). The reason is that the Gaussian rule is based on the zeros of the Legendre polynomials which are also used for expanding the phase function.

Before showing the matrix form of the discrete ordinate equations we must comment on the important consequence the number of streams (number of quadrature angles) $2N$ chosen has for the number of terms $2M$ in the Legendre polynomial expansion of the phase function, and thus for the number of azimuthal intensity components I^m (which is also $2M$ as discussed above). In (10) D^m is a polynomial of degree $2M - 1$. The maximum degree of a polynomial representing the intensity I^m is also $2M - 1$. The reason is that in the single-scattering approximation, when the intensity is expected to have the largest variability with polar angle, the intensity is proportional to the phase function (see Eqs. (94) and (95) later), which in the Legendre expansion is a polynomial of degree $2M - 1$. So, the product of D^m and I^m under the integral in (10) is a polynomial with a maximum degree $2(2M - 1)$. A $2N$-point Gauss quadrature is accurate for polynomials of up to degree $4N - 1$, so for an accurate evaluation of the integral in (10) we must choose the number of quadrature angles $2N$ such that $4N - 1 \geq 4M - 2$, from which it follows that $2N$ must at least equal $2M$. Because of this coupling between the number of streams and the number of Legendre coefficients in the phase function expansion in DISORT we set $2M = 2N$.

We now return to the discrete ordinate equations, and apply them to a vertically inhomogeneous atmosphere. In a vertically inhomogeneous medium, the coefficients D^m in (15) are functions of τ. The τ-dependence makes (14) a system of $2N$ coupled differential equations with nonconstant coefficients, for which there is no closed-form analytic solution. To obtain analytic solutions, DISORT assumes that the medium consists of L adjacent homogeneous layers. The single-scattering albedo and phase function are assumed to be constant within each layer, but allowed to vary from layer to layer (Fig. 2). Thus vertical variation of these is approximated by a step-function variation. The thermal source term is approximated by a polynomial variation in τ within each layer.

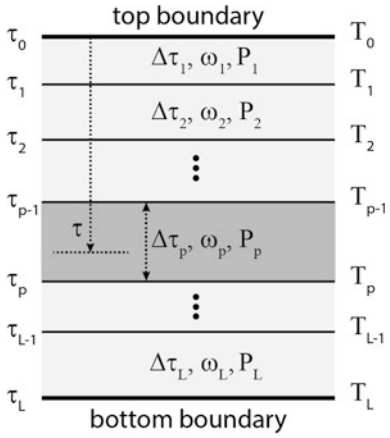


Fig. 2 Schematic illustration of a multilayered atmosphere (optical medium). Cumulative optical depth τ_p and temperature T_p ($p = 0, \ldots, L$) are defined at layer interfaces. Each layer is characterized by its layer-averaged single-scatter albedo ω, phase function P and optical thickness $\Delta\tau_p = \tau_p - \tau_{p-1}$

For illustration, we consider a single homogeneous layer $\tau_{p-1} \le \tau \le \tau_p$ but omit the p subscripts, which are implicit on all quantities. We also omit the τ arguments of D and S since they are, by assumption, independent of τ in a given layer.

To make the multi-stream case easier to understand, first we write (14) out for four streams ($N = 2$). (For two streams ($N = 1$) the equations are simpler, but the real structure is better seen in the four-stream case. For the two-stream case see Stamnes et al. 2000.) For $N = 2$ we obtain four coupled differential equations, one for each stream. Recalling that $\mu_{-i} = -\mu_i$, and $w_{-i} = w_i$ in the Gaussian quadrature rule applied, dividing both sides by μ's, and rearranging terms so they are ordered according to the directions μ_1, μ_2, $-\mu_1$, $-\mu_2$, and omitting the dependence of I^m and Q^m on τ, the corresponding equations are written as

$$\frac{\mathrm{d}I^m(\mu_1)}{\mathrm{d}\tau} = \frac{1}{\mu_1}\left\{\begin{array}{l} -[w_1 D^m(\mu_1,\mu_1) - 1]I^m(\mu_1) - w_2 D^m(\mu_1,\mu_2)I^m(\mu_2) \\ -w_1 D^m(\mu_1,-\mu_1)I^m(-\mu_1) - w_2 D^m(\mu_1,-\mu_2)I^m(-\mu_2) \end{array}\right\} - \frac{Q^m(\mu_1)}{\mu_1},$$

$$\frac{\mathrm{d}I^m(\mu_2)}{\mathrm{d}\tau} = \frac{1}{\mu_2}\left\{\begin{array}{l} -w_1 D^m(\mu_2,\mu_1)I^m(\mu_1) - [w_2 D^m(\mu_2,\mu_2) - 1]I^m(\mu_2) \\ -w_1 D^m(\mu_2,-\mu_1)I^m(-\mu_1) - w_2 D^m(\mu_2,-\mu_2)I^m(-\mu_2) \end{array}\right\} - \frac{Q^m(\mu_2)}{\mu_2},$$

$$\frac{\mathrm{d}I^m(-\mu_1)}{\mathrm{d}\tau} = \frac{1}{\mu_1}\left\{\begin{array}{l} w_1 D^m(-\mu_1,\mu_1)I^m(\mu_1) + w_2 D^m(-\mu_1,\mu_2)I^m(\mu_2) + \\ [w_1 D^m(-\mu_1,-\mu_1) - 1]I^m(-\mu_1) + w_2 D^m(-\mu_1,-\mu_2)I^m(-\mu_2) \end{array}\right\} - \frac{-Q^m(-\mu_1)}{\mu_1},$$

$$\frac{\mathrm{d}I^m(-\mu_2)}{\mathrm{d}\tau} = \frac{1}{\mu_2}\left\{\begin{array}{l} w_1 D^m(-\mu_2,\mu_1)I^m(\mu_1) + w_2 D^m(-\mu_2,\mu_2)I^m(\mu_2) + \\ w_1 D^m(-\mu_2,-\mu_1)I^m(-\mu_1) + [w_2 D^m(-\mu_2,-\mu_2) - 1]I^m(-\mu_2) \end{array}\right\} - \frac{-Q^m(-\mu_2)}{\mu_2}.$$

By defining the quantities

$$Q'^m(\pm\mu_i) = \pm\mu_i^{-1} Q^m(\pm\mu_i), \quad (i = 1, 2); \tag{17}$$

$$\begin{aligned}
\alpha_{11} &= \mu_1^{-1}[w_1 D^m(\mu_1,\mu_1) - 1] = \mu_1^{-1}[w_1 D^m(-\mu_1,-\mu_1) - 1],\\
\alpha_{22} &= \mu_2^{-1}[w_2 D^m(\mu_2,\mu_2) - 1] = \mu_2^{-1}[w_2 D^m(-\mu_2,-\mu_2) - 1],\\
\alpha_{12} &= \mu_1^{-1} w_2 D^m(\mu_1,\mu_2) = \mu_1^{-1} w_2 D^m(-\mu_1,-\mu_2),\\
\alpha_{21} &= \mu_2^{-1} w_1 D^m(\mu_2,\mu_1) = \mu_2^{-1} w_1 D^m(-\mu_2,-\mu_1),
\end{aligned} \tag{18}$$

$$\begin{aligned}
\beta_{11} &= \mu_1^{-1} w_1 D^m(\mu_1,-\mu_1) = \mu_1^{-1} w_1 D^m(-\mu_1,\mu_1),\\
\beta_{22} &= \mu_2^{-1} w_2 D^m(\mu_2,-\mu_2) = \mu_2^{-1} w_2 D^m(-\mu_2,\mu_2),\\
\beta_{12} &= \mu_1^{-1} w_2 D^m(\mu_1,-\mu_2) = \mu_1^{-1} w_2 D^m(-\mu_1,\mu_2),\\
\beta_{21} &= \mu_2^{-1} w_1 D^m(\mu_2,-\mu_1) = \mu_2^{-1} w_1 D^m(-\mu_2,\mu_1),
\end{aligned} \tag{19}$$

the four equations are written in matrix form as follows:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{bmatrix} I^m(\tau,\mu_1)\\ I^m(\tau,\mu_2)\\ I^m(\tau,-\mu_1)\\ I^m(\tau,-\mu_2)\end{bmatrix} &= \begin{bmatrix} -\alpha_{11} & -\alpha_{12} & -\beta_{11} & -\beta_{12}\\ -\alpha_{21} & -\alpha_{22} & -\beta_{21} & -\beta_{22}\\ \beta_{11} & \beta_{12} & \alpha_{11} & \alpha_{12}\\ \beta_{21} & \beta_{22} & \alpha_{21} & \alpha_{22}\end{bmatrix}\begin{bmatrix} I^m(\tau,\mu_1)\\ I^m(\tau,\mu_2)\\ I^m(\tau,-\mu_1)\\ I^m(\tau,-\mu_2)\end{bmatrix}\\
&\quad - \begin{bmatrix} Q'^m(\tau,\mu_1)\\ Q'^m(\tau,\mu_2)\\ Q'^m(\tau,-\mu_1)\\ Q'^m(\tau,-\mu_2)\end{bmatrix},
\end{aligned} \tag{20}$$

where we now include the τ dependence to emphasize that I^m and Q'^m generally vary within the homogeneous layer.

In (18) and (19) we took advantage of two symmetry properties of D^m: $D^m(\mu,\mu') = D^m(-\mu,-\mu')$ and $D^m(\mu,-\mu') = D^m(-\mu,\mu')$, which follow from (11) and from the fact that $\Lambda_\ell^m(-\mu) = (-1)^{\ell+m}\Lambda_\ell^m(\mu)$.

By introducing the vectors

$$\mathbf{I}^{\pm} = \{I^m(\tau,\pm\mu_i)\}, \quad \mathbf{Q}'^{\pm} = \{Q'(\tau,\pm\mu_i)\}, \quad (i=1,2), \tag{21}$$

and the matrices

$$\boldsymbol{\alpha} = \begin{bmatrix}\alpha_{11} & \alpha_{12}\\ \alpha_{21} & \alpha_{22}\end{bmatrix} \quad \text{and} \quad \boldsymbol{\beta} = \begin{bmatrix}\beta_{11} & \beta_{12}\\ \beta_{21} & \beta_{22}\end{bmatrix},$$

Equation (20) can be written in a more compact form as

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{bmatrix}\mathbf{I}^+\\ \mathbf{I}^-\end{bmatrix} = \begin{bmatrix}-\boldsymbol{\alpha} & -\boldsymbol{\beta}\\ \boldsymbol{\beta} & \boldsymbol{\alpha}\end{bmatrix}\begin{bmatrix}\mathbf{I}^+\\ \mathbf{I}^-\end{bmatrix} - \begin{bmatrix}\mathbf{Q}'^+\\ \mathbf{Q}'^-\end{bmatrix}. \tag{22}$$

From (18) and (19) it is clear that the matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ can be written as

$$\begin{aligned}
\boldsymbol{\alpha} &= \begin{bmatrix} \mu_1^{-1} & 0 \\ 0 & \mu_2^{-1} \end{bmatrix} \left\{ \begin{bmatrix} D(\mu_1,\mu_1) & D(\mu_1,\mu_2) \\ D(\mu_2,\mu_1) & D(\mu_2,\mu_2) \end{bmatrix} \begin{bmatrix} w_1 & 0 \\ 0 & w_2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} \\
&= \begin{bmatrix} \mu_1^{-1} & 0 \\ 0 & \mu_2^{-1} \end{bmatrix} \left\{ \begin{bmatrix} D(-\mu_1,-\mu_1) & D(-\mu_1,-\mu_2) \\ D(-\mu_2,-\mu_1) & D(-\mu_2,-\mu_2) \end{bmatrix} \begin{bmatrix} w_1 & 0 \\ 0 & w_2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right\} \\
&= \mathbf{M^{-1}}\{\mathbf{D}^+\,\mathbf{W} - \mathbf{1}\}, \\
\boldsymbol{\beta} &= \begin{bmatrix} \mu_1^{-1} & 0 \\ 0 & \mu_2^{-1} \end{bmatrix} \left\{ \begin{bmatrix} D(\mu_1,-\mu_1) & D(\mu_1,-\mu_2) \\ D(\mu_2,-\mu_1) & D(\mu_2,-\mu_2) \end{bmatrix} \begin{bmatrix} w_1 & 0 \\ 0 & w_2 \end{bmatrix} \right\} \\
&= \begin{bmatrix} \mu_1^{-1} & 0 \\ 0 & \mu_2^{-1} \end{bmatrix} \left\{ \begin{bmatrix} D(-\mu_1,\mu_1) & D(-\mu_1,\mu_2) \\ D(-\mu_2,\mu_1) & D(-\mu_2,\mu_2) \end{bmatrix} \begin{bmatrix} w_1 & 0 \\ 0 & w_2 \end{bmatrix} \right\} \\
&= \mathbf{M^{-1}}\{\mathbf{D}^-\mathbf{W}\},
\end{aligned}$$

where; $\mathbf{M} = \{\mu_i\delta_{ij}\}$; $\mathbf{W} = \{w_i\delta_{ij}\}$; $\mathbf{1} = \{\delta_{ij}\}$; $\mathbf{D}^+ = \{D^m(\mu_i,\mu_j)\} = \{D^m(-\mu_i,-\mu_j)\}$ and $\mathbf{D}^- = \{D^m(-\mu_i,\mu_j)\} = \{D^m(\mu_i,-\mu_j)\}$ $i,j = 1,2$.

It should now be relatively straightforward to generalize to $2N$ streams, so that (14) may be written in matrix form identically to (22):

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{bmatrix} \mathbf{I}^+ \\ \mathbf{I}^- \end{bmatrix} = \begin{bmatrix} -\boldsymbol{\alpha} & -\boldsymbol{\beta} \\ \boldsymbol{\beta} & \boldsymbol{\alpha} \end{bmatrix} \begin{bmatrix} \mathbf{I}^+ \\ \mathbf{I}^- \end{bmatrix} - \begin{bmatrix} \mathbf{Q}'^+ \\ \mathbf{Q}'^- \end{bmatrix}, \tag{23}$$

where

$$\begin{aligned}
\mathbf{I}^\pm &= \{I^m(\tau,\pm\mu_i)\}, \quad i = 1,\ldots,N, \\
\mathbf{Q}'^\pm &= \mathbf{M^{-1}}\mathbf{Q}^\pm, \\
\mathbf{Q}^\pm &= \{Q^m(\tau,\pm\mu_i)\}, \quad i = 1,\ldots,N, \\
\mathbf{M} &= \{\mu_i\delta_{ij}\}, \quad i,j = 1,\ldots,N,
\end{aligned} \tag{24}$$

$$\begin{aligned}
\boldsymbol{\alpha} &= \mathbf{M^{-1}}\{\mathbf{D}^+\mathbf{W} - \mathbf{1}\}, \\
\boldsymbol{\beta} &= \mathbf{M^{-1}}\{\mathbf{D}^-\,\mathbf{W}\}, \\
\mathbf{W} &= \{w_i\delta_{ij}\}, \quad i,j = 1,\ldots,N, \\
\mathbf{1} &= \{\delta_{ij}\}, \quad i,j = 1,\ldots,N, \\
\mathbf{D}^+ &= \{D^m(\mu_i,\mu_j)\} = \{D^m(-\mu_i,-\mu_j)\}, \quad i,j = 1,\ldots,N, \\
\mathbf{D}^- &= \{D^m(-\mu_i,\mu_j)\} = \{D^m(\mu_i,-\mu_j)\}, \quad i,j = 1,\ldots,N,
\end{aligned} \tag{25}$$

and where $\mathbf{1}$ is the identity matrix.

The special structure of the $(2N \times 2N)$ matrix

$$\begin{bmatrix} -\boldsymbol{\alpha} & -\boldsymbol{\beta} \\ \boldsymbol{\beta} & \boldsymbol{\alpha} \end{bmatrix}$$

can be traced to the reciprocity principle, which for single scattering is due to the fact that the phase function depends only on the scattering angle Θ (Chandrasekhar 1960, Chap. 4, Eq. 29). This special structure is also a consequence of having chosen a quadrature rule satisfying $\mu_{-i} = -\mu_i$ and $w_{-i} = w_i$. As we shall see later, this special structure leads to eigensolutions with eigenvalues occurring in positive/negative pairs, which allows a reduction in the order of the resulting algebraic eigenvalue problem by a factor of two. Since computation time for eigensolution algorithms grows roughly as the cube of the matrix dimension this reduction of order decreases the computational burden by a factor eight.

Stamnes (1986) notes that matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ may be interpreted as layer transmission and reflection operators, respectively. This interpretation can be seen, for example, by noticing that $\boldsymbol{\alpha}$ denotes quantities in which the D's are for the same directions (same sign of μ's), while $\boldsymbol{\beta}$ represents D's for opposite (upward and downward) directions (opposite sign of μ's).

4 Discrete Ordinate Approximation—Solution

Equation (14) is linear in the intensity, so traditionally, it is solved by combining two separate solutions: the "homogeneous" and "particular" solutions. The homogeneous solution must satisfy (14) with no source term, but it does not have to satisfy the boundary conditions; it typically contains arbitrary constants. The particular solution is a solution with the source terms included and no arbitrary constants, but again not required to satisfy the boundary conditions. The general solution is the sum of the homogeneous and particular solutions, which is additionally required to satisfy the boundary conditions. The boundary conditions are satisfied by solving linear algebraic equations for the arbitrary constants in the homogeneous solution.

Thus the whole process breaks up into three distinct parts: finding the homogeneous solution, finding the particular solution, and satisfying the boundary conditions.

4.1 Homogeneous Solution

Equation (23) is a system of $2N$ coupled, ordinary differential equations with constant coefficients. These coupled equations are linear and they can be uncoupled.

It is traditional to seek solutions to the homogeneous version ($\mathbf{Q}' = 0$) of (23) in the form

$$\mathbf{I}^{\pm} = \mathbf{G}^{\pm} e^{-k\tau}, \quad \mathbf{G}^{\pm} = \{G(\pm\mu_i)\} \quad (i = 1, \ldots, N). \tag{26}$$

Inserting this form into (23) we find

$$\begin{bmatrix} \boldsymbol{\alpha} & \boldsymbol{\beta} \\ -\boldsymbol{\beta} & -\boldsymbol{\alpha} \end{bmatrix} \begin{bmatrix} \mathbf{G}^{+} \\ \mathbf{G}^{-} \end{bmatrix} = k \begin{bmatrix} \mathbf{G}^{+} \\ \mathbf{G}^{-} \end{bmatrix}. \tag{27}$$

Equation (27) is a standard algebraic eigenvalue problem of order $2N \times 2N$; its solution determines the eigenvalues k and the eigenvectors $\mathbf{G}^{\pm}$. Because of the special structure of the matrix in (27), the eigenvalues occur in positive/negative pairs and the order of this algebraic eigenvalue problem may be reduced by a factor of two as follows (Stamnes and Swanson 1981). Expanding (27) we have

$$\begin{aligned} \boldsymbol{\alpha}\mathbf{G}^{+} + \boldsymbol{\beta}\mathbf{G}^{-} &= k\mathbf{G}^{+}, \\ -\boldsymbol{\beta}\mathbf{G}^{+} - \boldsymbol{\alpha}\mathbf{G}^{-} &= k\mathbf{G}^{-}. \end{aligned} \tag{28}$$

Adding and subtracting these two equations, we find

$$(\boldsymbol{\alpha} - \boldsymbol{\beta})(\mathbf{G}^{+} - \mathbf{G}^{-}) = k(\mathbf{G}^{+} + \mathbf{G}^{-}), \tag{29}$$

$$(\boldsymbol{\alpha} + \boldsymbol{\beta})(\mathbf{G}^{+} + \mathbf{G}^{-}) = k(\mathbf{G}^{+} - \mathbf{G}^{-}). \tag{30}$$

Expressing $(\mathbf{G}^{+} + \mathbf{G}^{-})$ from (29) and substituting the result into (30) yields

$$(\boldsymbol{\alpha} + \boldsymbol{\beta})(\boldsymbol{\alpha} - \boldsymbol{\beta})(\mathbf{G}^{+} - \mathbf{G}^{-}) = k^2(\mathbf{G}^{+} - \mathbf{G}^{-}). \tag{31}$$

Equation (31) is now an eigenvalue problem of order $N \times N$. A solution of (31) yields eigenvectors $\mathbf{G}^{+} - \mathbf{G}^{-}$ and the corresponding eigenvalues k^2. Since $(\boldsymbol{\alpha} + \boldsymbol{\beta})$ and $(\boldsymbol{\alpha} - \boldsymbol{\beta})$ are $N \times N$ matrixes there are two sets of eigenvalues, thus a total of $2N$ eigenvalues occurring in positive/negative pairs ($k = \pm\sqrt{k^2}$).

Substituting the solution $\mathbf{G}^{+} - \mathbf{G}^{-}$ into (29) we find $\mathbf{G}^{+} + \mathbf{G}^{-}$ (assuming $k \neq 0$)

$$\mathbf{G}^{+} + \mathbf{G}^{-} = \frac{1}{k}(\boldsymbol{\alpha} - \boldsymbol{\beta})(\mathbf{G}^{+} - \mathbf{G}^{-}). \tag{32}$$

To recover the eigenvectors $\mathbf{G}^{+}$ and $\mathbf{G}^{-}$ corresponding to the positive (k) and negative ($k = -k$) eigenvalues from their sum and difference we add and subtract the above solutions:

$$\begin{aligned} \mathbf{G}^{+} &= \frac{1}{2}[(\mathbf{G}^{+} + \mathbf{G}^{-}) + (\mathbf{G}^{+} - \mathbf{G}^{-})], \\ \mathbf{G}^{-} &= \frac{1}{2}[(\mathbf{G}^{+} + \mathbf{G}^{-}) - (\mathbf{G}^{+} - \mathbf{G}^{-})]. \end{aligned} \tag{33}$$

We now have a complete set of eigenvectors and we may write down the homogeneous solution as a linear combination of the solutions for different eigenvalues as follows:

$$I(\tau,\mu_i)=\sum_{\substack{j=-N\\ j\neq 0}}^{N} C_j\,G_j(\mu_i)\mathrm{e}^{-k_j\tau},\quad (i=-N,\ldots,-1,1,\ldots,N), \tag{34}$$

where the C_j are $2N$ constants of integration.

4.2 Particular Solution

To obtain the particular solution we consider the inhomogeneous term separately for the beam and the thermal sources, and then combine the solutions. For beam sources $Q^{(\mathrm{beam})}(\tau,\mu)=X_0(\mu)\mathrm{e}^{-\tau/\mu_0}$ and it is easily verified that a particular solution of (14) is (omitting m-superscript)

$$I(\tau,\mu_i)=Z_0(\mu_i)\,\mathrm{e}^{-\tau/\mu_0}, \tag{35}$$

where the Z_0 are determined by the following system of linear algebraic equations

$$\sum_{\substack{j=-N\\ j\neq 0}}^{N}\left[\left(1+\frac{\mu_j}{\mu_0}\right)\delta_{ij}-w_jD(\mu_i,\mu_j)\right]Z_0(\mu_j)=X_0(\mu_i),\quad (i=\pm1,\ldots,\pm N). \tag{36}$$

Equation (36) is a system of equations for the $2N$ unknowns, Z_0, which can readily be solved using standard linear equation solvers. Godsalve (1996) suggested that the order of the matrix inversion required to get Z_0 can be reduced, and this should improve computational efficiency. In DISORT, the reduction of order of the system is accomplished in a process similar to that used in the homogeneous solution (Lin et al. 2015). For this purpose, we rewrite (36) separately for the upward and downward directions in vector form

$$\left(\mathbf{1}+\frac{1}{\mu_0}\mathbf{M}-\mathbf{D}^{+}\mathbf{W}\right)\mathbf{Z_0^{+}}-\mathbf{D}^{-}\mathbf{W}\mathbf{Z_0^{-}}=\mathbf{X_0^{+}},\quad (i=+1,\ldots,+N), \tag{37}$$

$$\left(\mathbf{1}-\frac{1}{\mu_0}\mathbf{M}-\mathbf{D}^{+}\mathbf{W}\right)\mathbf{Z_0^{-}}-\mathbf{D}^{-}\mathbf{W}\mathbf{Z_0^{+}}=\mathbf{X_0^{-}},\quad (i=-1,\ldots,-N), \tag{38}$$

where $\mathbf{Z_0^{\pm}}=[Z_0(\pm\mu_1),Z_0(\pm\mu_2),\ldots Z_0(\pm\mu_N)]^T,\mathbf{X_0^{\pm}}=[X_0(\pm\mu_1),X_0(\pm\mu_2),\ldots X_0(\pm\mu_N)]^T$ and the other quantities have already been defined. Adding and subtracting (37) and (38) and rearranging terms we find

$$\mathbf{Z_0^+} - \mathbf{Z_0^-} = \mu_0\left[\mathbf{M^{-1}}\{\mathbf{D^+W} - \mathbf{1}\} + \mathbf{M^{-1}D^-W}\right]\left(\mathbf{Z_0^+} + \mathbf{Z_0^-}\right) + \mu_0\mathbf{M^{-1}}\left(\mathbf{X_0^+} + \mathbf{X_0^-}\right),$$
$$\mathbf{Z_0^+} + \mathbf{Z_0^-} = \mu_0\left[\mathbf{M^{-1}}\{\mathbf{D^+W} - \mathbf{1}\} - \mathbf{M^{-1}D^-W}\right]\left(\mathbf{Z_0^+} - \mathbf{Z_0^-}\right) + \mu_0\mathbf{M^{-1}}\left(\mathbf{X_0^+} - \mathbf{X_0^-}\right).$$

The first and the second terms in the square brackets are the matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ introduced earlier, so we write the above equations as

$$\mathbf{Z_0^+} - \mathbf{Z_0^-} = \mu_0[\boldsymbol{\alpha} + \boldsymbol{\beta}]\left(\mathbf{Z_0^+} + \mathbf{Z_0^-}\right) + \mu_0\mathbf{M^{-1}}\left(\mathbf{X_0^+} + \mathbf{X_0^-}\right), \tag{39}$$

$$\mathbf{Z_0^+} + \mathbf{Z_0^-} = \mu_0[\boldsymbol{\alpha} - \boldsymbol{\beta}]\left(\mathbf{Z_0^+} - \mathbf{Z_0^-}\right) + \mu_0\mathbf{M^{-1}}\left(\mathbf{X_0^+} - \mathbf{X_0^-}\right). \tag{40}$$

Substituting (40) into (39) we find

$$\left[\mathbf{1} - \mu_0^2(\boldsymbol{\alpha} + \boldsymbol{\beta})(\boldsymbol{\alpha} - \boldsymbol{\beta})\right]\left(\mathbf{Z_0^+} - \mathbf{Z_0^-}\right) = \mu_0^2\left[(\boldsymbol{\alpha} + \boldsymbol{\beta})\,\mathbf{M^{-1}}\left(\mathbf{X_0^+} - \mathbf{X_0^-}\right) + \frac{1}{\mu_0}\mathbf{M^{-1}}\left(\mathbf{X_0^+} + \mathbf{X_0^-}\right)\right], \tag{41}$$

which we solve for $\mathbf{Z_0^+} - \mathbf{Z_0^-}$. DISORT solves (41) using robust linear equation solving software, which protects against ill-conditioning automatically. With $\mathbf{Z_0^+} - \mathbf{Z_0^-}$ known we find $\mathbf{Z_0^+} + \mathbf{Z_0^-}$ from (40).

We recover $\mathbf{Z_0^+}$ and $\mathbf{Z_0^-}$ by adding and subtracting the vectors $\mathbf{Z_0^+} + \mathbf{Z_0^-}$ and $\mathbf{Z_0^+} - \mathbf{Z_0^-}$ as

$$\begin{aligned}\mathbf{Z_0^+} &= \frac{1}{2}\left[\left(\mathbf{Z_0^+} + \mathbf{Z_0^-}\right) + \left(\mathbf{Z_0^+} - \mathbf{Z_0^-}\right)\right],\\ \mathbf{Z_0^-} &= \frac{1}{2}\left[\left(\mathbf{Z_0^+} + \mathbf{Z_0^-}\right) - \left(\mathbf{Z_0^+} - \mathbf{Z_0^-}\right)\right].\end{aligned} \tag{42}$$

For thermal sources the emitted radiation is isotropic, so that

$$Q^0(\tau) = (1 - \omega)B[T(\tau)],$$
$$Q^m(\tau) = 0, \quad (m > 0).$$

DISORT assumes that the Planck function B varies linearly in optical depth across each layer, $B[T(\tau)] = b_0 + b_1\tau$ (Wiscombe 1976). The two coefficients b_0, b_1 are chosen to match B at the top and bottom layer boundaries τ_t and τ_b, where the temperatures are known

$$B_t \equiv B[T(\tau_t)] = b_0 + b_1\tau_t,$$
$$B_b \equiv B[T(\tau_b)] = b_0 + b_1\tau_b.$$

Since the thermal source term assumes the form

$$Q^{(\text{thermal})}(\tau) = (1-\omega)\,(b_0 + b_1\tau), \tag{43}$$

it suggests trying a particular solution of the form

$$I(\tau, \mu_i) = Y_0(\mu_i) + Y_1(\mu_i)\,\tau. \tag{44}$$

Substituting (44) into (14) and equating coefficients of like powers in τ, we find

$$\begin{aligned} \sum_{\substack{j=-N \\ j\neq 0}}^{N} \left[\delta_{ij} - w_j D^0(\mu_i, \mu_j)\right] Y_1(\mu_j) &= (1-\omega) b_1, \\ \sum_{\substack{j=-N \\ j\neq 0}}^{N} \left[\delta_{ij} - w_j D^0(\mu_i, \mu_j)\right] Y_0(\mu_j) &= (1-\omega) b_0 + \mu_i\, Y_1(\mu_i). \end{aligned} \tag{45}$$

Equations (45) are a system of linear algebraic equations determining $Y_0(\mu_i)$ and $Y_1(\mu_i)$.

4.3 General Solution

The general solution to (14) consists of a linear combination of all the homogeneous solutions, plus the particular solutions for beam and thermal emission sources (omitting m-superscript):

$$I(\tau, \mu_i) = \sum_{\substack{j=-N \\ j\neq 0}}^{N} C_j G_j(\mu_i)\, \mathrm{e}^{-k_j \tau} + Z_0(\mu_i)\, \mathrm{e}^{-\tau/\mu_0} + \delta_{m0}[Y_0(\mu_i) + Y_1(\mu_i)\,\tau]. \tag{46}$$

The k_j and $G_j(\mu_i)$ for $j \neq 0$ are the eigenvalues and eigenvectors, the μ_i the cosines of the quadrature angles, and the C_j the constants of integration to be determined by the boundary and layer continuity conditions.

4.4 Intensities at Arbitrary Angles

Equation (46) yields intensities at the $2N$ quadrature points. However, we may want intensities at directions that do not coincide with the quadrature points. For these directions the intensities are calculated from the formal solution of (9). For a slab of optical thickness τ_L, this solution is

$$
\begin{aligned}
I(\tau, +\mu) &= I(\tau_L, +\mu)\, \mathrm{e}^{-(\tau_L-\tau)/\mu} + \int_{\tau}^{\tau_L} S(t, +\mu)\, \mathrm{e}^{-(t-\tau)/\mu} \frac{\mathrm{d}t}{\mu}, \\
I(\tau, -\mu) &= I(0, -\mu)\, \mathrm{e}^{-\tau/\mu} + \int_{0}^{\tau} S(t, -\mu)\, \mathrm{e}^{-(\tau-t)/\mu} \frac{\mathrm{d}t}{\mu},
\end{aligned}
\tag{47}
$$

where $\mu > 0$, and we have again omitted the m-superscript. From these equations it is clear that if we know the source function $S(\tau, \pm\mu)$, we can find the intensity at arbitrary angles by integrating S. We now use the discrete ordinate solutions to derive explicit expressions for the source function. As we will see these expressions can be integrated analytically.

Using (12) with (43) for $Q^{(\text{thermal})}$ the source function (10) can be written as

$$
S(\tau, \mu) = \sum_{\substack{i=-N \\ i\neq 0}}^{N} w_i\, D(\mu, \mu_i)\, I(\tau, \mu_i) + X_0(\mu)\, \mathrm{e}^{-\tau/\mu_0} + \delta_{m0}(1-\omega)\,(b_0 + b_1\tau). \tag{48}
$$

Substituting the general solution of (46) into (48), we find

$$
S(\tau, \mu) = \sum_{j=-N}^{N} C_j\, G_j(\mu)\, \mathrm{e}^{-k_j\tau} + \delta_{m0}\,[V_0(\mu) + V_1(\mu)\,\tau], \tag{49}
$$

where

$$
C_0\, G_0(\mu) = Z_0(\mu) = \sum_{\substack{i=-N \\ i\neq 0}}^{N} w_i\, D(\mu, \mu_i)\, Z_0(\mu_i) + X_0(\mu), \quad k_0 = 1/\mu_0, \tag{50}
$$

$$
G_j(\mu) = \sum_{\substack{i=-N \\ i\neq 0}}^{N} w_i\, D(\mu, \mu_i)\, G_j(\mu_i) \quad \text{for } j \neq 0, \text{ and} \tag{51}
$$

$$
V_l(\mu) = \sum_{\substack{i=-N \\ i\neq 0}}^{N} w_i\, D^0(\mu, \mu_i)\, Y_l(\mu_i) + (1-\omega)\, b_l, \quad (l = 1, 0). \tag{52}
$$

In the equations above we have omitted the m-superscript, however, in (45) and (52) we have explicitly written D^0 as a reminder that thermal emission contributes only to the azimuth-independent component of the intensity.

We now present the discrete ordinate equations for computing the intensity at an arbitrary angle μ. However, before treating the general multilayer case, we first show the simpler, single-layer case ($L = 1$).

Substituting the source function (49) into the equations for the interpolated intensity (47) and calculating the simple integrals of exponentials analytically, we find that the downward and upward intensities become ($\mu > 0$):

$$
\begin{aligned}
I(\tau, +\mu) &= I(\tau_1, +\mu)\, \mathrm{e}^{-(\tau_1-\tau)/\mu} + \sum_{j=-N}^{N} C_j \frac{G_j(+\mu)}{1+k_j\mu} \left\{ \mathrm{e}^{-k_j\tau} - \mathrm{e}^{-\left[k_j\tau_1 + (\tau_1-\tau)/\mu\right]} \right\} \\
&\quad + \delta_{m0} \left\{ V_0(+\mu) \left[1 - \mathrm{e}^{-(\tau_1-\tau)/\mu} \right] + V_1(+\mu) \left[(\tau+\mu) - (\tau_1+\mu)\, \mathrm{e}^{-(\tau_1-\tau)/\mu} \right] \right\}, \\
I(\tau, -\mu) &= I(0, -\mu)\, \mathrm{e}^{-\tau/\mu} + \sum_{j=-N}^{N} C_j \frac{G_j(-\mu)}{1-k_j\mu} \left\{ \mathrm{e}^{-k_j\tau} - \mathrm{e}^{-\tau/\mu} \right\} \\
&\quad + \delta_{m0} \left\{ V_0(-\mu) \left[1 - \mathrm{e}^{-\tau/\mu} \right] + V_1(-\mu) \left[(\tau-\mu) + \mu\, \mathrm{e}^{-\tau/\mu} \right] \right\}.
\end{aligned}
\tag{53}
$$

In a multilayered medium the integral in (47) is evaluated by integrating layer by layer (cf. Stamnes 1982a) as follows ($\tau_{p-1} \le \tau \le \tau_p$ and $\mu > 0$):

$$
\begin{aligned}
&\int_{\tau}^{\tau_L} S(t, +\mu)\, \mathrm{e}^{-(t-\tau)/\mu} \frac{\mathrm{d}t}{\mu} = \int_{\tau}^{\tau_p} S_p(t, +\mu)\, \mathrm{e}^{-(t-\tau)/\mu} \frac{\mathrm{d}t}{\mu} + \sum_{n=p+1}^{L} \left\{ \int_{\tau_{n-1}}^{\tau_n} S_n(t, +\mu)\, \mathrm{e}^{-(t-\tau)/\mu} \frac{\mathrm{d}t}{\mu} \right\}, \\
&\int_{0}^{\tau} S(t, -\mu)\, \mathrm{e}^{-(\tau-t)/\mu} \frac{\mathrm{d}t}{\mu} = \sum_{n=1}^{p-1} \left\{ \int_{\tau_{n-1}}^{\tau_n} S_n(t, -\mu)\, \mathrm{e}^{-(\tau-t)/\mu} \frac{\mathrm{d}t}{\mu} \right\} + \int_{\tau_{p-1}}^{\tau} S_p(t, -\mu)\, \mathrm{e}^{-(\tau-t)/\mu} \frac{\mathrm{d}t}{\mu}.
\end{aligned}
\tag{54}
$$

Using (49) for $S_n(t, \mu)$ in (54), we find the intensities at τ in layer p

$$
\begin{aligned}
I_p(\tau, +\mu) &= I(\tau_L, +\mu)\, \mathrm{e}^{-(\tau_L-\tau)/\mu} \\
&\quad + \sum_{n=p}^{L} \left\{ \begin{array}{l} \displaystyle\sum_{j=-N}^{N} C_{jn} \frac{G_{jn}(+\mu)}{1+k_{jn}\,\mu} E_{jn}(\tau, +\mu) \\ \\ + \delta_{m0} [V_{0n}(+\mu) F_{0n}(\tau, +\mu) + V_{1n}(+\mu) F_{1n}(\tau, +\mu)] \end{array} \right\},
\end{aligned}
\tag{55}
$$

$$
\begin{aligned}
I_p(\tau, -\mu) &= I(0, -\mu)\, \mathrm{e}^{-\tau/\mu} \\
&\quad + \sum_{n=1}^{p} \left\{ \begin{array}{l} \displaystyle\sum_{j=-N}^{N} C_{jn} \frac{G_{jn}(-\mu)}{1-k_{jn}\,\mu} E_{jn}(\tau, -\mu) \\ \\ + \delta_{m0} \left[V_{0n}(-\mu)\ F_{0n}(\tau, -\mu) + V_{1n}(-\mu)\ F_{1n}(\tau, -\mu) \right] \end{array} \right\},
\end{aligned}
\tag{56}
$$

where

$$
\begin{aligned}
E_{jn}(\tau, +\mu) &= \exp\{-k_{jn}\tau_{n-1} - (\tau_{n-1} - \tau)/\mu\} - \exp\{-k_{jn}\tau_n - (\tau_n - \tau)/\mu\},\\
F_{0n}(\tau, +\mu) &= \exp\{-(\tau_{n-1} - \tau)/\mu\} - \exp\{-(\tau_n - \tau)/\mu\},\\
F_{1n}(\tau, +\mu) &= (\tau_{n-1} + \mu)\exp\{-(\tau_{n-1} - \tau)/\mu\} - (\tau_n + \mu)\exp\{-(\tau_n - \tau)/\mu\},
\end{aligned}
\tag{57}
$$

with τ_{n-1} replaced by τ for $n = p$, and

$$
\begin{aligned}
E_{jn}(\tau, -\mu) &= \exp\{-k_{jn}\tau_n - (\tau - \tau_n)/\mu\} - \exp\{-k_{jn}\tau_{n-1} - (\tau - \tau_{n-1})/\mu\},\\
F_{0n}(\tau, -\mu) &= \exp\{-(\tau - \tau_n)/\mu\} - \exp\{-(\tau - \tau_{n-1})/\mu\},\\
F_{1n}(\tau, -\mu) &= (\tau_n - \mu)\exp\{-(\tau - \tau_n)/\mu\} - (\tau_{n-1} - \mu)\exp\{-(\tau - \tau_{n-1})/\mu\},
\end{aligned}
\tag{58}
$$

with τ_n replaced by τ for $n = p$.

One can verify that substituting $\tau_{n-1} = \tau$, $\tau_n = \tau_L = \tau_1$ in (55), and $\tau_{n-1} = 0$, $\tau_n = \tau$ in (56) for a single layer, (55) and (56) reduce to the single-layer solutions (53). Also, when evaluated at the quadrature points for beam sources, (55) and (56) yield results identical to (46). In addition, (55) and (56) satisfy the boundary and continuity conditions for all values of μ, even though such conditions have been imposed only at the quadrature angles (Stamnes 1982a).

4.5 Boundary and Interface Conditions

The radiative transfer equation (2) must be solved subject to boundary conditions. In principle, the discrete ordinate method can handle quite general boundary conditions. However, the current implementation of the method in DISORT assumes that the medium is illuminated at the top boundary by a combination of known isotropic diffuse radiation and parallel beam treated as pseudosource. The bottom boundary is specified by a known bidirectional reflectance distribution function (BRDF) $\rho_d(\mu,\phi;-\mu',\phi')$ and thermal emission specified in terms of the Planck function B and directional emissivity $\varepsilon(\mu)$. Thus the boundary conditions in DISORT are specified as

$$
\begin{aligned}
I(\tau = 0, -\mu, \phi) &= I_{\text{top}}(\mu, \phi) = I_{\text{diffuse}} + F_0\,\delta(\mu - \mu_0)\,\delta(\phi - \phi_0),\\
I(\tau = \tau_L, +\mu, \phi) &= I_g(\mu, \phi) = \varepsilon(\mu)\,B(T_g) + \frac{1}{\pi}\mu_0 F_0\,\mathrm{e}^{-\tau_L/\mu_0}\,\rho_d(\mu, \phi; -\mu_0, \phi_0)\\
&\quad + \frac{1}{\pi}\int_0^{2\pi} \mathrm{d}\phi' \int_0^1 \rho_d(\mu, \phi; -\mu', \phi')\,I(\tau_L, -\mu', \phi')\,\mu'\mathrm{d}\mu'.
\end{aligned}
\tag{59}
$$

where I_{top} and I_g are the intensities incident at the top and bottom boundaries, respectively; and τ_L is the total optical depth of the entire medium. F_0 is the flux of the incident beam normal to the direction of incidence at the top boundary.

In DISORT we assume that the lower boundary is totally opaque and calculate the emissivity from the bidirectional reflectivity using Kirchhoff's law $\varepsilon(\mu) + \frac{1}{\pi}\int_0^{2\pi} \mathrm{d}\phi' \int_0^1 \rho_d(\mu, \phi; -\mu', \phi')\, \mu'\, \mathrm{d}\mu' = 1$. We also assume that the surface has no preferred direction. This means that ρ_d depends only on the difference $\phi - \phi'$ between the azimuthal angles of the incident and reflected radiation, and that the surface BRDF is symmetric about the principal plane. This symmetry allows us to expand the surface bidirectional reflectivity in a Fourier series of $2M$ terms just like we did for the intensity (8). Because of the assumed symmetry about the principal plane the Fourier expansion only contains cosine terms:

$$\rho_d(\mu, \phi; -\mu', \phi') = \rho_d(\mu, -\mu'; \phi - \phi') = \sum_{m=0}^{2M-1} \rho_d^m(\mu, -\mu') \cos m(\phi - \phi'), \quad (60)$$

where the expansion coefficients $\rho_d^m(\mu, -\mu')$ are computed from

$$\rho_d^m(\mu, -\mu') = (2 - \delta_{m0})\, \frac{1}{\pi} \int_0^{\pi} \rho_d(\mu, -\mu'; \phi - \phi') \cos m(\phi - \phi')\, \mathrm{d}(\phi - \phi').$$

We discuss how these coefficients are computed in DISORT in Sect. 7.4.

Substituting (60) into the second equation of (59) and using the Fourier expansion of the intensity (8) together with the quadratured form of the integral over μ', we find that each Fourier component must satisfy the bottom boundary condition

$$\begin{aligned} I^m(\tau_L, +\mu) = I_g^m(\mu) \equiv \delta_{m0}\, \varepsilon(\mu) B\big(T_g\big) + \frac{1}{\pi} \mu_0 F_0\, \mathrm{e}^{-\tau_L/\mu_0}\, \rho_d^m(\mu, -\mu_o) \\ + (1 + \delta_{m0}) \sum_{j=1}^{N} w_j \mu_j\, \rho_d^m\big(\mu, -\mu_j\big) I^m\big(\tau_L, -\mu_j\big). \end{aligned} \quad (61)$$

In a multilayered medium we must also require the intensity to be continuous across layer interfaces (Stamnes and Conklin 1984). Thus, (14) must satisfy boundary and continuity conditions as follows:

$$\begin{aligned} I_1^m(0, -\mu_i) &= I_{\text{top}}^m(\mu_i), \quad (i = 1, \ldots, N); \\ I_p^m\big(\tau_p, \mu_i\big) &= I_{p+1}^m\big(\tau_p, \mu_i\big), \quad (p = 1, \ldots, L; \quad i = \pm 1, \ldots, N); \\ I_L^m(\tau_L, +\mu_i) &= I_g^m(\mu_i), \quad (i = 1, \ldots, N); \end{aligned} \quad (62)$$

where the last equation is the same as (61) with μ replaced by μ_i.

We note that the first and the last equations of (62) introduce a fundamental distinction between downward ($-\mu$) and upward ($+\mu$) directions over which integrations should be done separately. The Double-Gauss rule adopted in DISORT readily satisfies this requirement.

For the discussion of boundary conditions, we write the discrete ordinate solution for the pth layer in the following form (omitting m-superscript)

$$I_p(\tau,\mu_i)=\sum_{j=1}^{N}\left\{C_{jp}G_{jp}(\mu_i)\mathrm{e}^{-k_{jp}\tau}+C_{-jp}G_{-jp}(\mu_i)\mathrm{e}^{+k_{jp}\tau}\right\}+U_p(\tau,\mu_i), \tag{63}$$

where $k_{jp} > 0$ and $k_{-jp} = -k_{jp}$, and the sum contains the homogeneous solution involving the unknown coefficients C_{jp} to be determined, and U_p is the particular solution given by (46):

$$U_p(\tau,\mu_i)=Z_0(\mu_i)\,\mathrm{e}^{-\tau/\mu_0}+\delta_{m0}[Y_0(\mu_i)+Y_1(\mu_i)\,\tau]. \tag{64}$$

Inserting (63) into (62) leads to

$$\sum_{j=1}^{N}\left\{C_{j1}G_{j1}(-\mu_i)+C_{-j1}G_{-j1}(-\mu_i)\right\}=I_{top}(-\mu_i)-U_1(0,-\mu_i),\quad \{i=1,\ldots,N\}; \tag{65}$$

$$\begin{aligned}\sum_{j=1}^{N}\Bigg\{&C_{jp}G_{jp}(\mu_i)\,\mathrm{e}^{-k_{jp}\tau_p}+C_{-jp}G_{-jp}(\mu_i)\,\mathrm{e}^{k_{jp}\tau_p}-\\&\left[C_{j,p+1}G_{j,p+1}(\mu_i)\,\mathrm{e}^{-k_{j,p+1}\tau_p}+C_{-j,p+1}G_{-j,p+1}(\mu_i)\,\mathrm{e}^{k_{j,p+1}\tau_p}\right]\Bigg\}\\&=U_{p+1}(\tau_p,\mu_i)-U_p(\tau_p,\mu_i),\quad \{p=1,\ldots,L-1;\ i=\pm1,\ldots,\pm N\};\end{aligned} \tag{66}$$

$$\sum_{j=1}^{N}\left\{\begin{matrix}C_{jL}\,r_{jL}(\mu_i)\,G_{jL}(\mu_i)\,\mathrm{e}^{-k_{jL}\tau_L}\\+C_{-jL}\,r_{-jL}(\mu_i)\,G_{-jL}(\mu_i)\,\mathrm{e}^{k_{jL}\tau_L}\end{matrix}\right\}=\Gamma(\tau_L,\mu_i),\quad \{i=1,\ldots,N\}; \tag{67}$$

where

$$r_{jL}(\mu_i)=1-(1+\delta_{m0})\sum_{n=1}^{N}\rho_d(\mu_i,-\mu_n)\,w_n\,\mu_n\,G_{jL}(-\mu_n)/G_{jL}(\mu_i), \tag{68}$$

$$\begin{aligned}\Gamma(\tau_L,\mu_i)=\delta_{m0}\,\varepsilon(\mu_i)\,B(T_g)-U_L(\tau_L,+\mu_i)+\frac{1}{\pi}\mu_0\,F_0\,\mathrm{e}^{-\tau_L/\mu_0}\rho_d(\mu_i,-\mu_0)\\+(1+\delta_{m0})\sum_{j=1}^{N}\rho_d(\mu_i,-\mu_j)\,w_j\,\mu_j\,U_L(\tau_L,-\mu_j).\end{aligned} \tag{69}$$

Equations (65)–(67) constitute a $(2N \times L) \times (2N \times L)$ system of linear algebraic equations from which the $2N \times L$ unknown coefficients C_{jp} ($j = \pm1, \ldots, \pm N$;

$p = 1, \ldots, L$) must be determined. The numerical solution of this set of equations is discussed in Sect. 7.2.

It is important to note that while the equations presented so far provide a mathematically correct, complete solution of the radiative transfer within the discrete ordinate approximation, Eqs. (65)–(67) are intrinsically ill-conditioned. Dealing with this ill-conditioning is absolutely crucial for the solutions to work. DISORT eliminates the ill-conditioning by a simple scaling transformation discussed in Sect. 4.6 below.

4.6 Scaling Transformation

Determination of the coefficients C_{jp} requires the solution of the system of linear algebraic Eqs. (65)–(67). These equations contain exponentials of the product $k_{jp}\tau_p$ that, because $k_j > 0$ by convention, is positive and becomes large for large τ_p. This situation leads to a matrix of coefficients having very small and very large values at the same time, and results in numerical ill-conditioning. To avoid the numerical ill-conditioning, it is necessary to remove the positive exponentials in (65)–(67). DISORT avoids the ill-conditioning by the scaling transformation developed by Stamnes and Conklin (1984).

If we write C_{+jp} and C_{-jp} as

$$C_{+jp} = \hat{C}_{+jp} e^{k_{jp}\tau_{p-1}} \quad \text{and} \quad C_{-jp} = \hat{C}_{-jp} e^{-k_{jp}\tau_p}, \tag{70}$$

insert them into (65)–(67) and solve for the $\hat{C}$'s instead of the C's, we find that all the exponential terms in the coefficient matrix have the form

$$\exp\left(-k_{jp}(\tau_p - \tau_{p-1})\right),$$

and because $k_{jp} > 0$ and $\tau_p > \tau_{p-1}$ all terms have negative arguments. This avoids numerical ill-conditioning and makes the solution for finding the $\hat{C}$'s unconditionally stable for arbitrary layer thicknesses. Since the particular solutions (64) do not have positive exponentials, the scaling transformation is only applied to the homogeneous solutions.

As an example, we demonstrate how the scaling works for the simplest case, the two-stream case ($N = 1$) for a one-layer medium ($L = 1$). In this simple case, (65)–(67) reduce to

$$\begin{aligned} C_{1,1}\, G_{1,1}(-\mu_1) + C_{-1,1}\, G_{-1,1}(-\mu_1) &= (\mathrm{RHS})_1, \\ r_{1,1}\, C_{1,1}\, G_{1,1}(+\mu_1)\, e^{-k_1\tau_1} + r_{-1,1}\, C_{-1,1}\, G_{-1,1}(+\mu_1)\, e^{k_1\tau_1} &= (\mathrm{RHS})_2, \end{aligned}$$

where "RHS" stands for "right-hand side". (Since for the two-stream case the only value j takes is 1, we have dropped the subscript j from k but left p because we shall

look at the two-layer case later.) Writing the left-hand sides in matrix form (and using the shorthand notation G^- for $-\mu$ and G^+ for $+\mu$) we have

$$\begin{bmatrix} G^-_{1,1} & G^-_{-1,1} \\ r_{1,1}G^+_{1,1}\mathrm{e}^{-k_1\tau_1} & r_{-1,1}G^+_{-1,1}\mathrm{e}^{k_1\tau_1} \end{bmatrix} \begin{bmatrix} C_{1,1} \\ C_{-1,1} \end{bmatrix}.$$

This matrix is ill-conditioned when $k_1\tau_1$ becomes large due to the positive exponential term $\exp(k_1\tau_1)$. Replacing the C's with their scaled values $C_{1,1} = \hat{C}_{1,1}\,\mathrm{e}^{k_1\tau_0}$ and $C_{-1,1} = \hat{C}_{-1,1}\,\mathrm{e}^{-k_1\tau_1}$ and carrying out the multiplications by the exponentials (keeping in mind that $\tau_0 = 0$) we find:

$$\begin{bmatrix} G^-_{1,1} & G^-_{-1,1}\mathrm{e}^{-k_1\tau_1} \\ r_{1,1}G^+_{1,1}\mathrm{e}^{-k_1\tau_1} & r_{-1,1}G^+_{-1,1} \end{bmatrix} \begin{bmatrix} \hat{C}_{1,1} \\ \hat{C}_{-1,1} \end{bmatrix},$$

where the positive exponential has now been removed. In the limit of large values of $k_1\tau_1$ this matrix becomes

$$\begin{bmatrix} G^-_{1,1} & 0 \\ 0 & r_{-1,1}G^+_{-1,1} \end{bmatrix},$$

which shows that the ill-conditioning problem has been entirely eliminated.

For the two-stream ($N = 1$), two-layer medium ($L = 2$) case (65)–(67) reduce to:

$$\begin{aligned}
&C_{1,1}G_{1,1}(-\mu_1) + C_{-1,1}G_{-1,1}(-\mu_1) = (\mathrm{RHS})_1,\\
&C_{1,1}G_{1,1}(-\mu_1)\,\mathrm{e}^{-k_1\tau_1} + C_{-1,1}G_{-1,1}(-\mu_1)\,\mathrm{e}^{k_1\tau_1}\\
&\quad - C_{1,2}G_{1,2}(-\mu_1)\,\mathrm{e}^{-k_2\tau_1} - C_{-1,2}G_{-1,2}(-\mu_1)\,\mathrm{e}^{k_2\tau_1} = (\mathrm{RHS})_2,\\
&C_{1,1}G_{1,1}(+\mu_1)\,\mathrm{e}^{-k_1\tau_1} + C_{-1,1}G_{-1,1}(+\mu_1)\,\mathrm{e}^{k_1\tau_1}\\
&\quad - C_{1,2}G_{1,2}(+\mu_1)\,\mathrm{e}^{-k_2\tau_1} - C_{-1,2}G_{-1,2}(+\mu_1)\,\mathrm{e}^{k_2\tau_1} = (\mathrm{RHS})_3,\\
&r_{1,2}C_{1,2}G_{1,2}(+\mu_1)\,\mathrm{e}^{-k_2\tau_2} + r_{-1,2}C_{-1,2}G_{-1,2}(+\mu_1)\,\mathrm{e}^{k_2\tau_2} = (\mathrm{RHS})_4.
\end{aligned}$$

Writing the left-hand sides of the above equations in matrix form we get:

$$\begin{bmatrix} G^-_{1,1} & G^-_{-1,1} & 0 & 0 \\ G^-_{1,1}\mathrm{e}^{-k_1\tau_1} & G^-_{-1,1}\mathrm{e}^{k_1\tau_1} & -G^-_{1,2}\mathrm{e}^{-k_2\tau_1} & -G^-_{-1,2}\mathrm{e}^{k_2\tau_1} \\ G^+_{1,1}\mathrm{e}^{-k_1\tau_1} & G^+_{-1,1}\mathrm{e}^{k_1\tau_1} & -G^+_{1,2}\mathrm{e}^{-k_2\tau_1} & -G^+_{-1,2}\mathrm{e}^{k_2\tau_1} \\ 0 & 0 & r_{1,2}\,G^+_{1,2}\mathrm{e}^{-k_2\tau_2} & r_{-1,2}\,G^+_{-1.2}\mathrm{e}^{k_2\tau_2} \end{bmatrix} \begin{bmatrix} C_{1,1} \\ C_{-1,1} \\ C_{1,2} \\ C_{-1,2} \end{bmatrix}.$$

Introducing the scaling transformation; that is, replacing the vector C with its scaled version $[\hat{C}_{1,1}\mathrm{e}^{k_1\tau_0},\ \hat{C}_{-1,1}\mathrm{e}^{-k_1\tau_1},\ \hat{C}_{1,2}\mathrm{e}^{k_2\tau_1},\ \hat{C}_{-1,2}\mathrm{e}^{-k_2\tau_2}]^T$ and carrying out the multiplications, we obtain

$$\begin{bmatrix} G^{-}_{1,1} & G^{-}_{-1,1}\mathrm{e}^{-k_1\tau_1} & 0 & 0 \\ G^{-}_{1,1}\mathrm{e}^{-k_1\tau_1} & G^{-}_{-1,1} & -G^{-}_{1,2} & -G^{-}_{-1,2}\mathrm{e}^{-k_2(\tau_2-\tau_1)} \\ G^{+}_{1,1}\mathrm{e}^{-k_1\tau_1} & G^{+}_{-1,1} & -G^{+}_{1,2} & -G^{+}_{-1,2}\mathrm{e}^{-k_2(\tau_2-\tau_1)} \\ 0 & 0 & r_{1,2}G^{+}_{1,2}\mathrm{e}^{-k_2(\tau_2-\tau_1)} & r_{-1,2}G^{+}_{-1,2} \end{bmatrix} \begin{bmatrix} \hat{C}_{1,1} \\ \hat{C}_{-1,1} \\ \hat{C}_{1,2} \\ \hat{C}_{-1,2} \end{bmatrix}.$$

Again, the transformation removes all positive exponentials as expected. The resulting matrix is well-conditioned for arbitrary layer optical thicknesses, and in the limit of large values of $k_1\tau_1$ and $k_2(\tau_2 - \tau_1)$ it reduces to

$$\begin{bmatrix} G^{-}_{1,1} & 0 & 0 & 0 \\ 0 & G^{-}_{-1,1} & -G^{-}_{1,2} & 0 \\ 0 & G^{+}_{-1,1} & -G^{+}_{1,2} & 0 \\ 0 & 0 & 0 & r_{-1,2}G^{+}_{-1,2} \end{bmatrix},$$

which again shows that the ill-conditioning problem has been entirely eliminated.

We now deal with the solutions for the general $2N$-stream, L-layer case. Since, as noted above, the scaling does not affect the particular solution below we write the solutions only for the homogeneous part.

Introducing the scaling into the homogeneous solution for the intensity at the quadrature (computational) angles in layer p (63) the transformed solution is written as

$$I_p(\tau, \mu_i) = \sum_{j=1}^{N} \left\{ \hat{C}_{jp} G_{jp}(\mu_i)\, \mathrm{e}^{-k_{jp}\left(\tau - \tau_{p-1}\right)} + \hat{C}_{-jp} G_{-jp}(\mu_i)\, \mathrm{e}^{-k_{jp}\left(\tau_p - \tau\right)} \right\}. \tag{71}$$

Since $k_{jp} > 0$ and $\tau_{p-1} \le \tau \le \tau_p$, all exponentials in (71) have negative arguments.

Writing the summation over j separately for the downward and upward directions (using $k_{-jp} = -k_{jp}$) in the homogeneous part of the solution for arbitrary upward directions (55) we have

$$I_p(\tau, +\mu) = \sum_{n=p}^{L} \sum_{j=1}^{N} \left\{ C_{-jn} \frac{G_{-jn}(+\mu)}{1 - k_{jn}\mu} E_{-jn}(\tau, +\mu) + C_{+jn} \frac{G_{+jn}(+\mu)}{1 + k_{jn}\mu} E_{+jn}(\tau, +\mu) \right\}. \tag{72}$$

The scaling transformation (70) changes Eq. (72) to

$$I_p(\tau, +\mu) = \sum_{n=p}^{L} \sum_{j=1}^{N} \left\{ \hat{C}_{-jn} \frac{G_{-jn}(+\mu)}{1 - k_{jn}\mu} \hat{E}_{-jn}(\tau, +\mu) + \hat{C}_{+jn} \frac{G_{+jn}(+\mu)}{1 + k_{jn}\mu} \hat{E}_{+jn}(\tau, +\mu) \right\}, \tag{73}$$

where

$$\hat{E}_{-jn}(\tau, +\mu) \equiv E_{-jn}(\tau, +\mu)\, \mathrm{e}^{-k_{jn}\tau_n} = \exp\left[-\left(k_{jn}\Delta\tau_n + \delta\tau/\mu\right)\right] - \exp[-(\tau_n - \tau)/\mu], \tag{74}$$

with

$$\begin{cases} \Delta\tau_n = \tau_n - \tau_{n-1}, & \delta\tau = \tau_{n-1} - \tau, & \text{for } n > p; \\ \Delta\tau_p = \tau_p - \tau, & \delta\tau = 0, & \text{for } n = p; \end{cases}$$

$$\hat{E}_{+jn}(\tau, +\mu) \equiv E_{+jn}(\tau, +\mu)\, \mathrm{e}^{k_{jn}\tau_{n-1}} = \exp\left[-\frac{\tau_{n-1} - \tau}{\mu}\right] - \exp\left\{-k_{jn}(\tau_n - \tau_{n-1}) - \frac{\tau_n - \tau}{\mu}\right\}, \tag{75}$$

for $n > p$ and

$$\hat{E}_{+jp}(\tau, +\mu) = \exp\left[-k_{jp}\left(\tau - \tau_{p-1}\right)\right] - \exp\left\{-k_{jp}\left(\tau_p - \tau_{p-1}\right) - \frac{\tau_p - \tau}{\mu}\right\}. \tag{76}$$

Since $k_{jn} > 0$ for $n = p + 1, p + 2, \ldots, L$ and $\tau_L > \cdots > \tau_{n=p+1} > \tau_{n-1=p} > \tau$ and also $k_{jp} > 0$ and $\tau_{p-1} < \tau < \tau_p$, all the exponentials in (74)–(76) have negative arguments.

Repeating the above for the homogeneous part of the solution for arbitrary downward directions (56), we find

$$I_p(\tau, -\mu) = \sum_{n=1}^{p} \sum_{j=1}^{N} \left\{ \hat{C}_{-jn} \frac{G_{-jn}(-\mu)}{1 + k_{jn}\mu} \hat{E}_{-jn}(\tau, -\mu) + \hat{C}_{+jn} \frac{G_{+jn}(-\mu)}{1 - k_{jn}\mu} \hat{E}_{+jn}(\tau, -\mu) \right\}, \tag{77}$$

where

$$\hat{E}_{+jn}(\tau, -\mu) \equiv E_{+jn}(\tau, -\mu)\, \mathrm{e}^{k_{jn}\tau_{n-1}} = \exp\left[-k_{jn}\Delta\tau_n - \frac{\delta\tau}{\mu}\right] - \exp\left[-\frac{\tau - \tau_{n-1}}{\mu}\right], \tag{78}$$

with

$$\begin{cases} \Delta\tau_n = \tau_n - \tau_{n-1}, & \delta\tau = \tau - \tau_n, & \text{for } n < p; \\ \Delta\tau_p = \tau - \tau_{p-1}, & \delta\tau = 0, & \text{for } n = p; \end{cases}$$

$$\begin{aligned}\hat{E}_{-jn}(\tau,-\mu) &\equiv E_{-jn}(\tau,-\mu)\mathrm{e}^{-k_{jn}\tau_n} \\ &= \exp\left[-\frac{\tau-\tau_n}{\mu}\right] - \exp\left\{-k_{jn}(\tau_n-\tau_{n-1}) - \frac{\tau-\tau_{n-1}}{\mu}\right\},\end{aligned} \tag{79}$$

for $n < p$ and

$$\hat{E}_{-jp}(\tau,-\mu) = \exp\left[-k_{jp}\left(\tau_p-\tau\right)\right] - \exp\left\{-k_{jp}\left(\tau_p-\tau_{p-1}\right) - \frac{\tau-\tau_{p-1}}{\mu}\right\}. \tag{80}$$

Again, we see that all exponentials involved have negative arguments since $k_{jn} > 0$ and $\tau > \tau_n > \tau_{n-1}$ for $n = 1, 2, \ldots, p-1$, and also $k_{jp} > 0$ and $\tau_{p-1} < \tau < \tau_p$. The negative arguments ensure that fatal overflow errors are avoided in the computations.

5 Flux, Flux Divergence, and Mean Intensity

Now that we have the intensities we can calculate some commonly used radiation quantities. Radiative fluxes that measure the total energy crossing a horizontal area per unit time in the upward (F^+) and downward (F^-) directions are defined as (Thomas and Stamnes 1999, Eq. 6.25):

$$\begin{aligned}F^+(\tau) &\equiv \int_0^{2\pi} \mathrm{d}\phi \int_0^1 \mu I(\tau,+\mu,\phi)\,\mathrm{d}\mu = 2\pi\int_0^1 \mu I^0(\tau,+\mu)\,\mathrm{d}\mu, \\ F^-(\tau) &\equiv \mu_0 F_0\,\mathrm{e}^{-\tau/\mu_0} + \int_0^{2\pi}\mathrm{d}\phi\int_0^1 \mu I(\tau,-\mu,\phi)\,\mathrm{d}\mu \\ &= \mu_0 F_0\,\mathrm{e}^{-\tau/\mu_0} + 2\pi\int_0^1 \mu I^0(\tau,-\mu)\,\mathrm{d}\mu.\end{aligned} \tag{81}$$

In these equations the second equalities are obtained by substituting the Fourier expansion of the intensity (8) for I and carrying out the integration over azimuth, which eliminates all but the $m = 0$ term of I^m.

Mean intensities measure the total intensity entering a volume from all sides. The only difference between them and the fluxes is that fluxes are cosine-weighted

averages of the intensity while mean intensities are unweighted averages. The hemispherical mean intensities are defined correspondingly as

$$\begin{aligned}\bar{I}^{+}(\tau) &\equiv \frac{1}{2\pi}\int_0^{2\pi} \mathrm{d}\phi \int_0^1 I(\tau, +\mu, \phi)\,\mathrm{d}\mu = \int_0^1 I^0(\tau, +\mu)\,\mathrm{d}\mu,\\ \bar{I}^{-}(\tau) &\equiv \frac{1}{2\pi}\left[F_0\,\mathrm{e}^{-\tau/\mu_0} + \int_0^{2\pi} \mathrm{d}\phi \int_0^1 I(\tau, -\mu, \phi)\,\mathrm{d}\mu\right]\\ &= \frac{1}{2\pi}F_0\,\mathrm{e}^{-\tau/\mu_0} + \int_0^1 I^0(\tau, -\mu)\,\mathrm{d}\mu.\end{aligned} \tag{82}$$

It can be seen from (81) and (82) that

$$I^0(\tau, \pm\mu) \equiv \frac{1}{2\pi}\int_0^{2\pi} I(\tau, \pm\mu, \phi)\,\mathrm{d}\phi, \tag{83}$$

which is the azimuthally averaged intensity. Note that in (81) and (82) we have also included the direct beam contribution in the downward flux F^- and in the downward mean intensity $\bar{I}^-$.

The net flux and mean intensity are then (Thomas and Stamnes 1999, Eq. 6.26)

$$\begin{aligned}F(\tau) &\equiv F^-(\tau) - F^+(\tau),\\ \bar{I}(\tau) &\equiv \frac{1}{2}[\bar{I}^-(\tau) + \bar{I}^+(\tau)].\end{aligned} \tag{84}$$

From the mean intensity the flux divergence, which is related to the radiative heating rate, can now be calculated as (Thomas and Stamnes 1999, Eq. 6.27):

$$\frac{\mathrm{d}F}{\mathrm{d}\tau} = 4\pi(1-\omega)\{\bar{I} - B[T(\tau)]\}. \tag{85}$$

DISORT uses Gaussian quadrature for the interval [0, 1] with abscissae μ_i and weights w_i, i = 1 to N. Thus, the DISORT approximations for flux and mean intensity are:

$$
\begin{aligned}
F^+(\tau) &= 2\pi \int_0^1 \mu I^0(\tau, +\mu)\,\mathrm{d}\mu \cong 2\pi \sum_{i=1}^{N} w_i \mu_i I^0(\tau, +\mu_i), \\
F^-(\tau) &= \mu_0 F_0\, \mathrm{e}^{-\tau/\mu_0} + 2\pi \int_0^1 \mu I^0(\tau, -\mu)\,\mathrm{d}\mu \\
&\cong \mu_0 F_0\, \mathrm{e}^{-\tau/\mu_0} + 2\pi \sum_{i=1}^{N} w_i \mu_i I^0(\tau, -\mu_i),
\end{aligned}
\tag{86}
$$

$$
\begin{aligned}
\bar{I}(\tau) &= \frac{1}{4\pi} F_0\, \mathrm{e}^{-\tau/\mu_0} + \frac{1}{2}\left[\int_0^1 I^0(\tau, +\mu)\,\mathrm{d}\mu + \int_0^1 I^0(\tau, -\mu)\,\mathrm{d}\mu\right] \\
&\cong \frac{1}{4\pi} F_0\, \mathrm{e}^{-\tau/\mu_0} + \frac{1}{2}\sum_{i=1}^{N} w_i \left[I^0(\tau, +\mu_i) + I^0(\tau, -\mu_i) \right].
\end{aligned}
\tag{87}
$$

These equations show that if we are only interested in fluxes and/or mean intensities, we need only the Fourier term ($m = 0$) I^0.

6 Optimizing Efficiency and Accuracy for Intensities

We have shown that the number of streams ($2N$) in the discrete ordinate approximation is tied to the number of moments in the Legendre expansion of the phase function P, which in turn determines the number of terms in the azimuthal series of the intensity (8).

There are four situations where the sum in (8) collapses to the $m = 0$ term: (1) only fluxes and/or mean intensities are needed (as we have already seen above); (2) there are no beam sources; (3) the beam source is at the zenith ($\mu_0 = 1$); (4) only zenith and/or nadir intensities are needed. We can understand these special cases in terms of their complete azimuthal symmetry about the vertical direction.

In all other cases (14) must be solved repeatedly for every m in the azimuthal series (8). Strongly forward-peaked, highly asymmetric phase functions require a very large number of terms (sometimes more than several hundred) in the expansion of P, and thus many terms in the intensity expansion. At the same time, for accurate intensities a similarly large number of streams ($2N$) is needed. These circumstances together can make the computation of intensities expensive.

6.1 δ–M Transformation

To achieve optimum computational efficiency and accuracy DISORT uses the δ–M transformation (Wiscombe 1977). The goal of the transformation is to reduce the length of the phase function expansion (5) from $2M$, which can be very large, 100s to 1000s, to $2N$, which might be as little as 4–8. The δ–M method achieves this reduction by separating the phase function P into the sum of a highly anisotropic phase function P'', approximated as a delta function in the forward direction, and a less anisotropic phase function P' which is expanded in a series of Legendre polynomials:

$$\begin{aligned} P(\tau, \cos\Theta) &= f\,P''(\tau, \cos\Theta) + (1-f)\,P'(\tau, \cos\Theta) \\ &\approx 2f\,\delta(1-\cos\Theta) + (1-f)\sum_{\ell=0}^{2N-1}(2\ell+1)\,g'_\ell(\tau)\,P_\ell(\cos\Theta). \end{aligned} \tag{88}$$

The δ–M Legendre expansion coefficients g'_ℓ are obtained from the Legendre coefficients g_ℓ of the phase function P by

$$g'_\ell(\tau) = \frac{g_\ell(\tau) - f}{1-f}, \quad (\ell = 0, \ldots, 2N-1); \tag{89}$$

and the separated fraction f is chosen by setting

$$f = g_{2N}(\tau). \tag{90}$$

To see why δ–M is an improvement let us consider the expansion coefficients in P and P'. In the ordinary phase function (P) expansion $g_\ell = 0$ for $\ell \geq 2N$. In P', g'_ℓ is also zero for $\ell \geq 2N$, but by (89, 90) this is equivalent to $g_\ell = g_{2N}$ for $\ell \geq 2N$, and this latter approximation should be better than $g_\ell = 0$ since typically the g_ℓ decrease only slowly with ℓ.

Substituting (88) into the radiative transfer equation (2), an identical form of the transfer equation is obtained, but with P', τ', ω' replacing P, τ, ω, respectively, where

$$\begin{aligned} d\tau' &= (1-\omega f)\,d\tau, \\ \omega' &= \frac{1-f}{1-\omega f}\,\omega. \end{aligned} \tag{91}$$

Application of the δ–M method artificially enhances the direct beam component of flux at the expense of the diffuse component, but the sum of the two is computed accurately. The true direct beam component is also accurately calculated as an exponential attenuation of the incident direct beam using the unscaled optical depth.

DISORT then recovers the accurate diffuse component by subtracting the true direct beam component from the δ–M scaled sum.

By reducing the length of the phase function expansion, and thus the number of streams and azimuthal intensity components required, the δ–M transformation decreases the computational burden. It has proven to be very useful for flux computations, but produces unacceptable error in intensity computations of practical size of streams ($2N < 10$–20), especially in the bright region close to the direct beam, called the aureole. The next section describes the method DISORT uses to alleviate this problem.

6.2 Correction of the Intensity Field

As discussed in the previous section, the δ–M transformation leads to errors for intensities. We illustrate the problem in Fig. 3, which shows the phase function P and its δ–M versions P' with $2N = 32$ and $2N = 8$. We use the Henyey–Greenstein (H–G) phase function with $g = 0.75$ in this illustration because it is analytical, and because its Legendre expansion coefficients can be readily obtained. The δ–M phase function $P'(2N = 32)$ reconstructed from $2N = 32$ Legendre moments reproduces the analytical P(H–G) quite well. Also, because f in this case is very small ($g_{32} = 10^{-4}$), the δ–M scaling has little impact. So we use $P'(2N = 32)$ and associated intensities calculated by DISORT as a reference for comparison below. We see that P' with $2N = 8$, $P'(2N = 8)$, is substantially less peaked than P in the forward direction ($\Theta = 0^{o}$), thus making $P'(2N = 8)$ less anisotropic as desired. However, this truncation makes representation of forward scattering less accurate. We also see that $P'(2N = 8)$ oscillates around the correct phase function.

The solid and dashed lines in Fig. 4 (top panels) show the intensities $I'(2N = 32)$ and $I'(2N = 8)$ corresponding to phase functions $P'(2N = 32)$ and $P'(2N = 8)$,

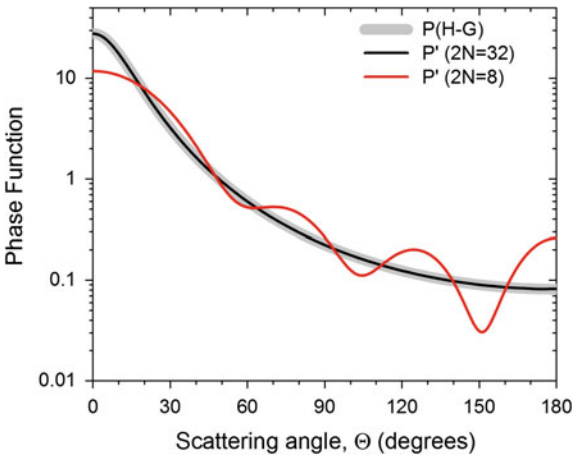


Fig. 3 Accurate phase function and its δ–M approximations with $2N = 8$ and with $2N = 32$ Legendre moments. Henyey–Greenstein phase function with asymmetry parameter $g = 0.75$ is used for illustration

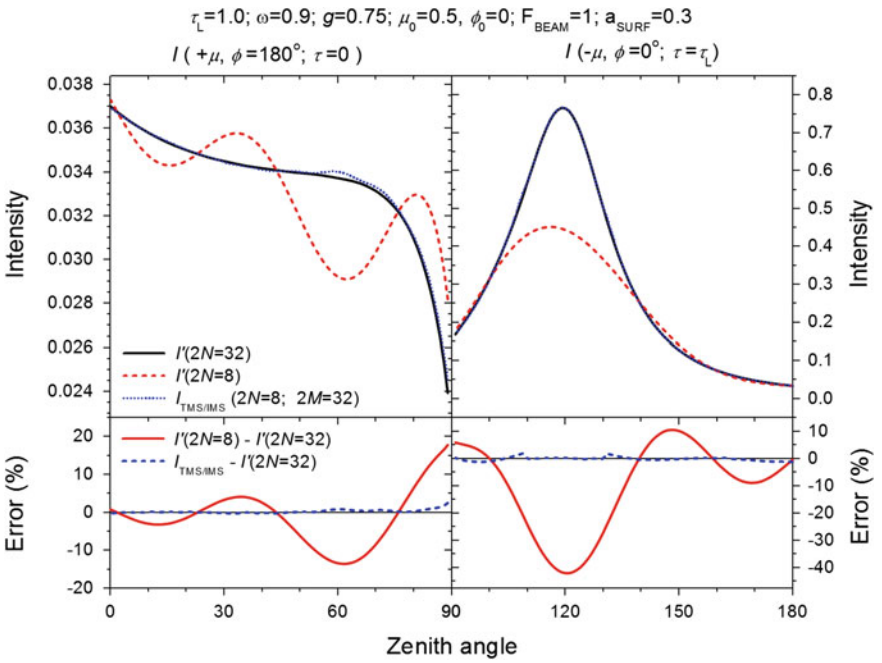


Fig. 4 *δ–M* scaled upward (*top-left*) and downward (*top-right*) intensities calculated with 32 [$I'(2N = 32)$] and 8 [$I'(2N = 8)$] streams corresponding to phase functions $P'(2N = 32)$ and $P'(2N = 8)$, respectively, for a slab of optical depth 1.0 and $\omega = 0.9$ bounded from below by a Lambertian reflecting surface with an albedo 0.3. Errors in 8-stream intensities relative to the accurate 32-stream intensities are plotted in the bottom panels. The 8-stream, 32-moment TMS/IMS-corrected intensities ($I_{\text{TMS/IMS}}$) and their error, again, relative to the 32-stream intensities are also shown

respectively, for a slab of optical depth 1.0 and $\omega = 0.9$ bounded from below by a Lambertian reflecting surface with an albedo 0.3. We see that, similar to $P'(2N = 8)$, the δ–M transformed intensity $I'(2N = 8)$ also oscillates around the reference (approximating the non-transformed) intensity. The oscillation is especially large for the downward intensity (right panel), but it is also significant for the upward intensities (left panel) as demonstrated by the relative difference plots (bottom panels).

Nakajima and Tanaka (1988) noticed that the oscillation in intensity resembles that in the phase function P'. They suggested that low-order of scattering, which is proportional to the phase function, contributed most to the oscillations, and proposed a correction, the so-called TMS and IMS methods, that would mitigate it. The idea is to subtract the less accurate, δ–M transformed, low-order scattering contributions from the δ–M transformed intensity I' and add back the "accurate", non-transformed, low-order scattered intensity. DISORT implements both the TMS and IMS methods.

The TMS method corrects the single-scattered intensity for using the δ–M phase function and the δ–M scaling of optical depth and single-scattering albedo, such that the TMS-corrected intensity I_{TMS} is

$$I_{\mathrm{TMS}} = I' - I'_{ss} + I_{ss}, \tag{92}$$

where I'_{ss} and I_{ss} are the δ–M transformed and accurate (no δ–M) single-scattered intensities, respectively. They are the solutions of the radiative transfer for single scattering shown below.

The single-scattering special case of the radiative transfer equation (2) is obtained by neglecting the multiple scattering source term in (3), which leads to:

$$\pm\mu \frac{dI_{ss}(\tau, \pm\mu, \phi)}{\mathrm{d}\tau} = I_{ss}(\tau, \pm\mu, \phi) - \frac{\omega(\tau)}{4\pi} P(\tau, +\mu, \phi; -\mu_0, \phi_0) F_0 \, \mathrm{e}^{-\tau/\mu_0}, \tag{93}$$

where the subscript "*ss*" refers to single scattering. Here the notation $+\mu$ indicates upward, while $-\mu$ indicates downward directions with $\mu > 0$ (note that (93) also ignores the thermal source term Q^{thermal}). Equation (93) can be readily solved by multiplying both sides by an integrating factor, $\exp(-\tau/\mu)$ for upward intensity and $\exp(+\tau/\mu)$ for downward intensity, and then by integrating layer by layer from τ_{L} to τ for upward intensity and from τ_0 to τ for downward intensity subject to lower and upper boundary conditions, respectively.

The solution for the upward intensity is

$$\begin{aligned} I_{ss}(\tau, +\mu, \phi) = {} & \frac{F_0}{4\pi(1+\mu/\mu_0)} \sum_{n=p}^{L} \omega_n P_n(+\mu, \phi; -\mu_0, \phi_0) \\ & \times \left\{ \exp\left(-\frac{\tau_{n-1} - \tau}{\mu} - \frac{\tau_{n-1}}{\mu_0} \right) - \exp\left(-\frac{\tau_n - \tau}{\mu} - \frac{\tau_n}{\mu_0} \right) \right\}, \end{aligned} \tag{94}$$

with τ_{n-1} replaced by τ for $n = p$. The downward intensity for $\mu \neq \mu_0$ is

$$\begin{aligned} I_{ss}(\tau, -\mu, \phi) = {} & \frac{F_0}{4\pi\,(1-\mu/\mu_0)} \sum_{n=1}^{p} \omega_n \, P_n(-\mu, \phi; -\mu_0, \phi_0) \\ & \times \left\{ \exp\left(-\frac{\tau - \tau_n}{\mu} - \frac{\tau_n}{\mu_0} \right) - \exp\left(-\frac{\tau - \tau_{n-1}}{\mu} - \frac{\tau_{n-1}}{\mu_0} \right) \right\}, \end{aligned} \tag{95}$$

with τ_n replaced by τ for $n = p$. In the special case $\mu = \mu_0$ the solution of (93) leads to

$$I_{ss}(\tau, -\mu, \phi) = \frac{F_0}{4\pi\mu_0} \mathrm{e}^{-\tau/\mu_0} \sum_{n=1}^{p} \omega_n \, P_n(-\mu, \phi; \, -\mu_0, \phi_0) \, (\tau_n - \tau_{n-1}), \tag{96}$$

with τ_n replaced by τ for $n = p$. Since $\tau_L > \cdots > \tau_n > \tau_{n-1} \geq \tau$ in (94) and $\tau > \cdots > \tau_n > \tau_{n-1} \geq 0$ in (95), all exponentials have negative arguments.

We note that the derivation of (94)–(96) assumes $I(\tau_0) = 0$ and $I(\tau_L) = 0$. This assumption is technically correct since both are diffuse intensities and as such they have already been scattered at least once; thus they are not part of the single-scattering solutions.

Equations (94) and (95/96) are solved twice: once with the δ–M scaled single-scattering albedo and optical depth (91) and phase function P' calculated from the $2M$ δ–M Legendre expansion coefficients (89), and once with the unscaled single-scattering albedo, scaled optical depth, and the exact phase function P using all available Legendre moments, but scaled by $1/(1 - \omega f)$ because of the δ–M τ scaling. The first solution gives I'_{ss}, while the second provides I_{ss} in (92). Writing the statements above in functional forms using the schematic notation

$$\text{Intensity} = \mathbf{F}[\text{single-scattering albedo} \times \text{phase function, optical depth}]$$

we have

$$I'_{ss} = \mathbf{F}[\omega' P'(2M), \tau'] \quad \text{and} \quad I_{ss} = \mathbf{F}\left[\frac{\omega\, P(N_{\max})}{1 - \omega f}, \tau'\right], \tag{97}$$

where the "primed" variables denote δ–M scaled quantities, and the argument of P indicates the number of Legendre moments in the expansion.

We emphasize again that for the TMS correction to be effective, P must be calculated from all Legendre moments $N_{\max}$ needed to adequately represent the exact phase function, not only from the $2M$ (=$2N$) moments used by DISORT with a given number of streams ($2N$).

The TMS method substantially reduces the errors in intensity except in the solar aureole near zero degree scattering angle. Further reduction of the error in the aureole can be achieved by using the IMS method (Nakajima and Tanaka 1988).

The error remaining in the intensity after the TMS correction is the difference between the "true" intensity and the TMS-corrected intensity. In DISORT, the IMS intensity correction term ΔI_{IMS} is defined as the negative of this difference

$$\Delta I_{\text{IMS}} \equiv -\left(I_{\text{true}} - I_{\text{TMS}}\right), \tag{98}$$

so it will be subtracted from the TMS-corrected DISORT intensity when applied. Note that by its definition ΔI_{IMS} includes intensities scattered more than once.

Differentiating both sides of (98) and using the δ–M scaling relations (91) and the definition of I_{TMS} (92) in the result, we get

$$\begin{aligned} -\mu \frac{\mathrm{d}\{\Delta I_{\text{IMS}}\}}{\mathrm{d}\tau} &= -\mu\left[\frac{\mathrm{d}I_{\text{IMS}}}{\mathrm{d}\tau} - \frac{\mathrm{d}I_{\text{true}}}{\mathrm{d}\tau}\right] \\ &= \mu \frac{\mathrm{d}I_{\text{true}}}{\mathrm{d}\tau} - \mu[1 - \omega(\tau) f(\tau)] \frac{\mathrm{d}}{\mathrm{d}\tau'}\left\{ I' - I'_{ss} + I_{ss} \right\}. \end{aligned} \tag{99}$$

Substituting (2)–(4) into (99), and using the δ–M approximation for the phase function (88) and the equation for the single-scattering solution (93), and rearranging the terms we find

$$-\mu \frac{\mathrm{d}\{\Delta I_{\mathrm{IMS}}\}}{\mathrm{d}\tau} = \Delta I_{\mathrm{IMS}} - [Q_1 + Q_2 + Q_3] - \omega(\tau)\,[P * \Delta I_{\mathrm{IMS}}], \tag{100}$$

where the Q's are

$$\begin{aligned} Q_1 &= \omega(\tau) f(\tau) \left\{ I'_{\mathrm{mult}} - \left[P'' * I'_{\mathrm{mult}}\right] \right\}, \\ Q_2 &= \omega(\tau) f(\tau) \left\{ I_{ss} - [P'' * I_{ss}] \right\}, \\ Q_3 &= \frac{F_0\,\omega(\tau)}{4\pi} P\left(\tau, -\mu, \phi; -\mu_0, \phi_0\right) \left[\mathrm{e}^{-\tau'/\mu_0} - \mathrm{e}^{-\tau/\mu_0} \right] \\ &\quad - \omega(\tau)[1 - f(\tau)] \left[P' * \left(I_{ss} - I'_{ss}\right)\right]. \end{aligned} \tag{101}$$

Here $I'_{\mathrm{mult}} = I' - I'_{ss}$ is the δ–M intensity due to multiple scattering, and P'' is the δ–M residual phase function derived from (88):

$$P''(\tau, \cos\Theta) = \frac{1}{f}\{P(\tau, \cos\Theta) - (1 - f)P'(\tau, \cos\Theta)\}. \tag{102}$$

The notation $[P * I]$ is used to write, in a short form, the integral of the product of the indicated phase function and intensity

$$[P * I] \equiv \frac{1}{4\pi} \int_0^{2\pi} \mathrm{d}\phi' \int_{-1}^{1} P(\tau, \mu, \phi; \mu', \phi') I(\tau, \mu', \phi')\, \mathrm{d}\mu'. \tag{103}$$

In these equations the Q and $[P * I]$ all have arguments $(\tau, -\mu, \phi)$.Equation (101) is solved by making a series of approximations. The most important of these comes from the intent to apply the IMS correction only around the solar aureole region ($\mu = \mu_0$), as the method has originally been designed to improve the retrieval of atmospheric aerosol properties from the analysis of solar aureole data obtained by sun photometers (Nakajima 2010). In this region, the source function Q_1 and the angular integral term in (100) can be neglected due to their small contribution, leading to:

$$-\mu \frac{\mathrm{d}\{\Delta I_{\mathrm{IMS}}\}}{\mathrm{d}\tau} = \Delta I_{\mathrm{IMS}} - [Q_2 + Q_3]. \tag{104}$$

The intensities in Q_2 and Q_3 are all single-scattered intensities, and so $[P * I]$ leads to intensities scattered twice. Therefore, the solution of (104) shall provide correction for second-order scattering.

Because P'' is a delta function in the δ–M method,

$$\frac{1}{4\pi}\int_0^{2\pi} \mathrm{d}\phi' \int_{-1}^{1} P''(\tau,-\mu,\phi;\,\mu',\phi')P'(\tau,\mu',\phi';-\mu_0,\phi_0)\,\mathrm{d}\mu' = P'(\tau,-\mu,\phi;-\mu_0,\phi_0). \tag{105}$$

By substituting the solutions for the two single-scattered intensities the source term Q_2 becomes,

$$Q_2(\tau,-\mu,\phi) \approx \frac{F_0}{4\pi}\frac{[\omega(\tau)f(\tau)]^2}{1-\omega(\tau)f(\tau)}\left[P''(\tau,-\mu,\phi;-\mu_0,\phi_0) - \frac{1}{4\pi}\int_0^{2\pi}\mathrm{d}\phi' \int_{-1}^{1} P''(\tau,-\mu,\phi;\,\mu',\phi')\,P''(\tau,\mu',\phi';-\mu_0,\phi_0)\,\mathrm{d}\mu'\right] \times \varphi(\tau',-\mu_0,-\mu_0). \tag{106}$$

Here, the approximation is that the geometrical factor φ

$$\varphi(\tau,-\mu,-\mu_0) \equiv \frac{\mathrm{e}^{-\tau/\mu}}{\mu}\int_0^{\tau} \mathrm{e}^{(1/\mu-1/\mu_0)\,t}\mathrm{d}t, \tag{107}$$

is nearly independent of μ near $\mu = \mu_0$ and can be brought outside the angular integrals by setting $\mu = \mu_0$. By using (91), and considering only the first term ($n = 1$) of the Taylor series expansion, the exponential term in Q_3 can be approximated as

$$\mathrm{e}^{-\tau'/\mu_0} - \mathrm{e}^{-\tau/\mu_0} = \mathrm{e}^{-\tau'/\mu_0}\left[1-\mathrm{e}^{-\omega f\,\tau/\mu_0}\right] = -\mathrm{e}^{-\tau'/\mu_0}\sum_{n=1}^{\infty}\frac{1}{n!}\left[\frac{-\omega(\tau)f(\tau)\,\tau}{\mu_0}\right]^n \approx \frac{\omega(\tau)f(\tau)}{1-\omega(\tau)f(\tau)}\varphi(\tau',-\mu_0,-\mu_0). \tag{108}$$

Using (108) and following the pattern of (106), the source term Q_3 becomes

$$Q_3(\tau,-\mu,\phi) \approx \frac{F_0}{4\pi}\frac{\omega^2(\tau)f(\tau)}{1-\omega(\tau)f(\tau)}\left[P(\tau,-\mu,\phi;-\mu_0,\phi_0) - \frac{1-f(\tau)}{4\pi} \times \int_0^{2\pi}\mathrm{d}\phi' \int_{-1}^{1} P'(\tau,-\mu,\phi;\,\mu',\phi')\,P''(\tau,\mu',\phi';-\mu_0,\phi_0)\,\mathrm{d}\mu'\right]\varphi(\tau',-\mu_0,-\mu_0) = \frac{F_0}{4\pi}\frac{[\omega(\tau)f(\tau)]^2}{1-\omega(\tau)f(\tau)}P''(\tau,-\mu,\phi;-\mu_0,\phi_0)\varphi(\tau',-\mu_0,-\mu_0). \tag{109}$$

Since the IMS correction is applied only to the transmitted intensity around the solar aureole, the contribution to secondary scattered intensity from backward scattering is relatively small. Thus, for a vertically inhomogeneous, layered atmosphere layer-averaged optical properties of the medium should be sufficient. (104) then has an analytical solution of

$$\begin{aligned}\Delta I_{\mathrm{IMS}} \approx \frac{F_0}{4\pi}\frac{(\bar{\omega}\bar{f})^2}{1-\bar{\omega}\bar{f}}\left[2P''(-\mu,\phi;-\mu_0,\phi_0) - P''^2(-\mu,\phi;-\mu_0,\phi_0)\right] \\ \times \chi\left(\tau,-\mu,-\mu_0',-\mu_0'\right),\end{aligned} \tag{110}$$

where the square of P'' is defined as an operator:

$$P''^2(\tau,-\mu,\phi;-\mu_0,\phi_0) \equiv \frac{1}{4\pi}\int_0^{2\pi} \mathrm{d}\phi' \int_{-1}^{1} P''(\tau,-\mu,\phi;\mu',\phi')P''(\tau,\mu',\phi';-\mu_0,\phi_0)\,\mathrm{d}\mu', \tag{111}$$

and the mean optical properties of the atmosphere are

$$\begin{aligned}\bar{\omega} &= \sum_{n=1}^{p}\omega_n\tau_n \Big/ \sum_{n=1}^{p}\tau_n, \\ \bar{f} &= \sum_{n=1}^{p} f_n\omega_n\tau_n \Big/ \sum_{n=1}^{p}\omega_n\tau_n, \\ P''(\cos\Theta) &= \sum_{\ell=0}^{N_{\max}}(2\ell+1)\,\bar{g}_\ell\,P_\ell(\cos\Theta), \\ \bar{g}_\ell &= \sum_{n=1}^{p} g'_{\ell,n}\omega_n\tau_n \Big/ \sum_{n=1}^{p} f_n\omega_n\tau_n, \\ g'_{\ell,n} &\equiv \begin{cases} f_n & \ell \le 2N-1 \\ g_{\ell,n} & \ell > 2N-1. \end{cases}\end{aligned} \tag{112}$$

In (112) P'' is the residual phase function; it is not a delta function, it is only approximated as a delta function in the δ–M method. $N_{\max}$ is the number of terms necessary to converge the Legendre expansion of the exact phase function P.

The function χ in (110) is

$$\chi(\tau,-\mu,-\mu',-\mu'') = \frac{\mathrm{e}^{-\tau/\mu}}{\mu\,\mu'}\int_0^{\tau}\mathrm{e}^{t(1/\mu-1/\mu')}\mathrm{d}t \int_0^{t}\mathrm{e}^{t'(1/\mu-1/\mu'')}\mathrm{d}t', \tag{113}$$

which is obtained by using the transformation of geometrical factor

$$\varphi(\tau', -\mu_0, -\mu_0) = \varphi(\tau, -\mu_0', -\mu_0'), \quad \mu_0' \equiv \frac{\mu_0}{1 - \bar{\omega}\bar{f}}.$$

Carrying out the integration in (113) requires distinguishing five different cases for different combinations of μ, μ', $\mu'' > 0$:

$$\chi = \frac{1}{\mu\,\mu'} \begin{cases} \frac{\tau^2 e^{-\tau/\mu}}{2} & \mu' = \mu'',\ \mu = \mu' \\ \frac{1}{x_1}\left[\left(\tau - \frac{1}{x_1}\right)e^{-\tau/\mu'} + \frac{e^{-\tau/\mu}}{x_1}\right] & \mu' = \mu'',\ \mu \neq \mu' \\ \frac{1}{x_2}\left[\frac{e^{-\tau/\mu''} - e^{-\tau/\mu}}{x_2} - \tau\, e^{-\tau/\mu}\right] & \mu' \neq \mu'',\ \mu = \mu' \\ \frac{1}{x_1}\left[\frac{e^{-\tau/\mu'} - e^{-\tau/\mu}}{x_1} - \tau\, e^{-\tau/\mu}\right] & \mu' \neq \mu'',\ \mu = \mu'' \\ \frac{1}{x_2}\left[\frac{e^{-\tau/\mu''} - e^{-\tau/\mu}}{x_2} - \frac{e^{-\tau/\mu'} - e^{-\tau/\mu}}{x_1}\right] & \mu \neq \mu' \neq \mu'', \end{cases} \tag{114}$$

where

$$x_1 = \frac{1}{\mu} - \frac{1}{\mu'} \quad \text{and}\, x_2 = \frac{1}{\mu} - \frac{1}{\mu''}.$$

Substituting the definition (112) of the residual phase function P'' into (111), and expressing the phase functions in terms of the directions $(-\mu, \varphi)$ and $(-\mu_0, \varphi_0)$ by using the addition theorem for spherical harmonics (7), and then carrying out the integration leads to

$$P''^2(\tau, -\mu, \phi; -\mu_0, \phi_0) = \sum_{\ell=0}^{N_{\max}} (2\ell + 1)\, \bar{g}_\ell^2\, P_\ell(\cos\Theta). \tag{115}$$

Substituting (112) and (115) into (110), the IMS intensity correction term ΔI_{IMS} becomes

$$\Delta I_{\text{IMS}} \approx \frac{F_0}{4\pi} \frac{(\bar{\omega}\bar{f})^2}{1 - \bar{\omega}\bar{f}} \left[\sum_{\ell=0}^{N_{\max}} (2\ell + 1)\left(2\bar{g}_\ell - \bar{g}_\ell^2\right) P_\ell(\cos\Theta)\right] \chi\left(\tau, -\mu, -\mu_0', -\mu_0'\right). \tag{116}$$

DISORT applies the IMS correction for downward intensities within ±10° of the solar beam direction.

Figure 4 shows the corrected intensities $I_{\text{TMS/IMS}}$ (top panels) and the differences relative to the reference intensity (bottom panels). The oscillations seen in the δ–M transformed intensity $I'(2N = 8)$ are greatly reduced by the intensity correction for the entire range of polar angles. The corrected downward intensity (top-right panel), for example, is virtually indistinguishable from the reference intensity. The relative difference (bottom-right panel) hovers within ±2 % of the zero line; maximum

values are at 110° and 130° where the IMS correction is turned on and off, respectively.

We noted that the single-scattering solutions used in the TMS correction neglect the reflection from the lower boundary. Similarly, the IMS correction does not account for radiation reflected by the surface either. Lin et al. (2015) show, however, that adding a correction term, ΔI_{BRDF} to (92) due to reflection of the direct beam at the lower boundary

$$\begin{aligned}\Delta I_{\mathrm{BRDF}} = &\left[\rho_d(+\mu, -\mu_0; \phi - \phi_0) - \rho_{d,\mathrm{appr}}(+\mu, -\mu_0; \phi - \phi_0)\right] \\ &\times \frac{1}{\pi}\mu_0 F_0\, \mathrm{e}^{-\tau'_L/\mu_0}\mathrm{e}^{-(\tau'_L-\tau')/\mu},\end{aligned} \tag{117}$$

improves accuracy of $I(+\mu, \phi)$, especially around the "hot spot." Here ρ_d is the accurate bidirectional reflectance distribution function (known for example, from an analytical function) and $\rho_{d,\mathrm{appr}}$ is the approximate BRDF calculated from the finite $2M$ Fourier expansion terms in (60). Note that (117) corrects for the error caused by using a series expansion of ρ_d that is shorter than necessary for an accurate representation of the BRDF; this correction has nothing to do with using the δ–M method. Also note that for a Lambertian surface ΔI_{BRDF} vanishes since in this case $\rho_{d,\mathrm{appr}} = \rho_d$.

Combining all three corrections the final, corrected intensity in DISORT is

$$\begin{aligned}I_{\mathrm{corrected}} &= I_{\mathrm{TMS}} - \Delta I_{\mathrm{IMS}} + \Delta I_{\mathrm{BRDF}} \\ &= I' - I'_{ss} + I_{ss} - \Delta I_{\mathrm{IMS}} + \Delta I_{\mathrm{BRDF}}.\end{aligned} \tag{118}$$

DISORT applies the intensity corrections after all azimuthal components of the δ–M scaled intensity are computed and summed up in (8) to give I'. As such, the intensity correction is a post-processing step.

Because the Nakajima–Tanaka intensity corrections are not applied to I^m, the azimuthally averaged intensity I^0 is uncorrected; therefore DISORT does not return it in the output. We note, however, that in principle the intensity corrections could also be applied to the Fourier components of the intensity. The necessary formulas, for example, for the TMS and IMS corrections, can be derived as follows. Denoting the TMS and IMS-corrected intensity by $I_{\mathrm{TMS/IMS}}$ and the TMS component of the correction by $\Delta I_{TMS} \equiv I_{ss} - I'_{ss}$

$$I_{\mathrm{TMS/IMS}} = I' + \Delta I_{\mathrm{TMS}} - \Delta I_{\mathrm{IMS}}.$$

Expanding the Is and ΔIs into an $N_{\max}$-term Fourier cosine series in azimuth, we have

$$\sum_{m=0}^{N_{\max}} I^m(\tau, -\mu)\, \cos m(\phi - \phi_0) = \sum_{m=0}^{N_{\max}} \left[I^{*m}(\tau, -\mu) + \Delta I^m_{TMS}(\tau, -\mu) + \Delta I^m_{\mathrm{IMS}}(\tau, -\mu)\right] \cos m(\phi - \phi_0),$$

$$\sum_{m=0}^{N_{\max}} I^m(\tau, +\mu)\, \cos m(\phi - \phi_0) = \sum_{m=0}^{N_{\max}} \left[I^{*m}(\tau, +\mu) + \Delta I^m_{\mathrm{TMS}}(\tau, +\mu)\right] \cos m(\phi - \phi_0),$$

where

$$I^{*m} = \begin{cases} I'^{m} & 0 \le m \le 2N-1 \\ 0 & 2N \le m \le N_{\max} \end{cases}.$$

From these equations the mth Fourier components of the downward and upward intensity, respectively, are

$$I^m(\tau,-\mu) = \begin{cases} I'^m(\tau,-\mu) + \Delta I^m_{TMS}(\tau,-\mu) + \Delta I^m_{\mathrm{IMS}}(\tau,-\mu), & 0 \le m \le 2N-1, \\ \Delta I^m_{\mathrm{TMS}}(\tau,-\mu) + \Delta I^m_{\mathrm{IMS}}(\tau,-\mu), & 2N \le m \le N_{\max}, \end{cases}$$

$$I^m(\tau,+\mu) = \begin{cases} I'^m(\tau,+\mu) + \Delta I^m_{\mathrm{TMS}}(\tau,+\mu), & 0 \le m \le 2N-1, \\ \Delta I^m_{\mathrm{TMS}}(\tau,+\mu), & 2N \le m \le N_{\max}. \end{cases}$$

These equations simply state that the δ–M scaled intensities vanish for $m \ge 2N$, and that the IMS correction is only applied for the downward directions.

The mth Fourier component of the TMS correction term, $\Delta I^m_{\mathrm{TMS}} = I^m_{ss} - I'^m_{ss}$, requires the mth Fourier components of the single-scattering solutions obtained from the unscaled and the δ–M scaled phase functions. These are obtained from Eqs. (94)–(96) by expanding the intensity in a Fourier cosine series and the phase function in a series of $2N$ normalized associated Legendre polynomials:

$$I'^m_{ss}(\tau',-\mu) = \frac{F_0}{4\pi\,(1-\mu/\mu_0)}\,(2-\delta_{m0}) \sum_{n=1}^{p} \left\{ \sum_{\ell=m}^{2N-1} g'_{\ell,n}\,\Lambda^m_\ell(-\mu)\,\Lambda(-\mu_0) \right\} \times \left\{ \exp\left(-\frac{\tau'-\tau'_n}{\mu} - \frac{\tau'_n}{\mu_0}\right) - \exp\left(-\frac{\tau'-\tau'_{n-1}}{\mu} - \frac{\tau'_{n-1}}{\mu_0}\right) \right\},$$

$$I'^m_{ss}(\tau',-\mu=-\mu_0) = \frac{F_0}{4\pi\,(1-\mu/\mu_0)}\,\mathrm{e}^{-\tau/\mu_0}\,(2-\delta_{m0}) \times \sum_{n=1}^{p} \left\{ \sum_{\ell=m}^{2N-1} g'_{\ell,n}\,\Lambda^m_\ell(-\mu)\,\Lambda(-\mu_0) \right\} \left(\tau'_n - \tau'_{n-1}\right),$$

$$I'^m_{ss}(\tau',+\mu) = \frac{F_0}{4\pi\,(1+\mu/\mu_0)}\,(2-\delta_{m0}) \sum_{n=p}^{L} \left\{ \sum_{\ell=m}^{2N-1} g'_{\ell,n}\,\Lambda^m_\ell(+\mu)\,\Lambda(-\mu_0) \right\} \times \left\{ \exp\left(-\frac{\tau'_{n-1}-\tau'}{\mu} - \frac{\tau'_{n-1}}{\mu_0}\right) - \exp\left(-\frac{\tau'_n-\tau'}{\mu} - \frac{\tau'_n}{\mu_0}\right) \right\}.$$

I^m_{ss} is calculated similarly, except the summation over ℓ runs up to $N_{\max}$ instead of $2N-1$, and $g'_{\ell,n}$ is replaced by $g_{\ell,n}/(1-\omega_n f_n)$.

The Fourier components of the IMS correction term ΔI_{IMS} is obtained from (116). Here the intensity is expanded into an $N_{\max}$-term Fourier cosine series and the Legendre polynomial $P_\ell(\cos\Theta)$ is rewritten in terms of the normalized associated Legendre polynomials $\Lambda^m_\ell(\mu)$ leading to

$$\Delta I_{\mathrm{IMS}} \approx \frac{F_0}{4\pi} \frac{\left(\bar{\omega}\bar{f}\right)^2}{1-\bar{\omega}\bar{f}} \left(2-\delta_{0m}\right) \times \left[\sum_{\ell=0}^{N_{\max}} \left(2\ell+1\right)\left(2\bar{g}_\ell - \bar{g}_\ell^2\right) \Lambda_\ell^m(-\mu)\Lambda_\ell^m(-\mu_0)\right] \chi\left(\tau, -\mu, -\mu_0', -\mu_0'\right),$$

where $\bar{\omega}, \bar{f}$ and $\bar{g}_\ell$ are given in (112).

We also note here that Buras et al. (2011) have developed an alternative method to calculate secondary scattering. Instead of expanding the phase function in terms of Legendre polynomials, they use the scattering phase function directly as a function of scattering angle. They numerically integrate (111) in a way that enhances sampling of the phase function in the forward peak. The method can provide more accurate intensities, particularly for strongly forward-scattering phase functions, because it is not compromised by the finite expansion. It is also more efficient because for DISORT to match accuracy it must use a very large number of Legendre moments $N_{\max}$.

7 Numerical Considerations

7.1 Computation of Eigenvalues and Eigenfunctions

To find the eigenvalues and eigenvectors needed for the intensity in (34) we must solve (31). Since the matrix $\mathbf{A} = (\boldsymbol{\alpha} - \boldsymbol{\beta})(\boldsymbol{\alpha} + \boldsymbol{\beta})$ in (31) is real but asymmetric, one may adopt subroutines like those in EISPACK (cf. Cowell 1980) which utilize the double QR algorithm to find eigenvalues and eigenvectors (Stamnes and Swanson 1981).

We note that because all eigenvalues are real, as it was shown by Stamnes and Swanson (1981), it is possible to transform this algebraic eigenvalue problem into one involving symmetric matrices. Such attempts were made by Nakajima and Tanaka (1986) and by Stamnes et al. (1988b). Stamnes et al. (1988b) also evaluated the merits of the transformation to symmetric matrices, and concluded that the transformation and subsequent solution procedures involved matrix operations in which the effect of rounding error became significant, and there was no significant increase in speed from symmetrizing the eigenproblem, because of the overhead of extra matrix multiplications. The advantage of solving the algebraic eigenvalue problem involving the asymmetric matrix $\boldsymbol{A}$ is that only one matrix multiplication is necessary.

The double QR algorithm applies to a general real matrix (which typically has complex eigenvalues/eigenvectors) and thus uses complex arithmetic. Use of complex arithmetic wastes a large amount of computation since only real arithmetic is needed for the discrete ordinates application. Therefore, one of the coauthors (Tsay) developed a customized version of the eigenvalue problem solver. This

customized version, ASYMTX, is an adaptation of subroutine EIGRF in the IMSL (International Mathematics and Statistics Library) to use real instead of complex arithmetic. EIGRF is based primarily on EISPACK routines. The matrix is first balanced using the Parlett–Reinsch algorithm (Parlett and Reinsch 1969). Then the Martin–Wilkinson algorithm (Wilkinson 1965) is applied. Stamnes et al. (2000) show the details of how ASYMTX works. These are not repeated here.

The eigen-computation in DISORT is handled entirely by ASYMTX. It is vital that this subroutine be run in the highest possible precision—at least IEEE double precision, or 14 significant digits. The reason is that the eigenproblem becomes degenerate as the single-scatter albedo approaches unity (see Sect. 7.5), and in general because it is a sensitive computation requiring high precision.

7.2 Numerical Solution for the Constants of Integration

We have shown that application of the boundary conditions and continuity conditions at layer interfaces leads to (65)–(67). These equations constitute a $(2N \times L) \times (2N \times L)$ system of linear algebraic equations for the $2N \times L$ unknown coefficients C_{jp} ($j = \pm 1, \ldots, \pm N$; $p = 1, \ldots, L$). This set of equations is ill-conditioned for large total optical thickness or large individual layer thickness, so a scaling transformation (70) was introduced that led to (73) and (77) for the homogeneous part of the solution with $\hat{C}_{jp}$ replacing C_{jp}.

The coefficient matrix for this system of algebraic equations for determining $\hat{C}_{jp}$ is a banded matrix with $(6N - 1)$ diagonals (Stamnes and Conklin 1984). Numerical packages like LINPACK (Dongarra et al. 1979) and LAPACK (Anderson et al. 1999) include solvers specifically designed for banded matrices. A banded matrix of order I and band width J can be solved in a time proportional to IJ^2. In our case, $I = 2NL$ and $J = N$ implying that the run time will increase as LN^3, that is linearly with the number of layers and cubically with the number of streams. Versions of DISORT use either LINPACK or LAPACK (see Sect. 9).

7.3 Removable Singularities in the Intensities

Certain combinations of the directions and eigenvalues may lead to 0/0-type singularities in the intensities and in (55) and (56). These singularities are removed by applying L'Hospital's Rule. The discrete ordinate solutions for the pth layer (55), (56) contain the following expressions:

$$I_p(\tau, +\mu) \propto \frac{E_{jn}(\tau, +\mu)}{1 + k_{jn}\mu} \exp\left(k_{jn}\tau_n\right) \equiv f_1\left(\tau, k_{jn}\mu\right),$$
$$I_p(\tau, -\mu) \propto \frac{E_{jn}(\tau, -\mu)}{1 - k_{jn}\mu} \exp\left(k_{jn}\tau_{n-1}\right) \equiv f_2\left(\tau, k_{jn}\mu\right).$$

The E_{jn} are given in (57) and (58). The exponential factors on the right-hand sides come from applying the scaling transformation (70). Whenever $|k_{jn}\mu| = 1$, the upward and downward intensities have a 0/0-type singularity for the eigenvalues $k_{jn} < 0$ and $k_{jn} > 0$, respectively. Applying L'Hospital's Rule leads to

$$f_1\left(\tau, k_{jn}\mu = -1\right) = \frac{\tau_n - \tau_{n-1}}{\mu} \exp(-(\tau_n - \tau)/\mu),$$
$$f_2\left(\tau, k_{jn}\mu = +1\right) = \frac{\tau_n - \tau_{n-1}}{\mu} \exp(-(\tau - \tau_{n-1})/\mu).$$

The particular solution of the downward intensity for the beam source $I_p(\tau,-\mu)$ is proportional to

$$f(\tau, -\mu, \mu_0) \equiv \frac{E_{jn}(\tau, -\mu)}{1 - \mu/\mu_0}.$$

The singularity at $\mu = \mu_0$ is again avoided by applying L'Hospital's Rule leading to

$$f(\tau, -\mu = \mu_0, \mu_0) = \frac{\tau_n - \tau_{n-1}}{\mu_0} \exp(-\tau/\mu_0).$$

Another potential singularity in the particular solution occurs when the square of the cosine of the solar zenith angle is too close to the reciprocal of an eigenvalue, and so the first factor on the left-hand side of (41) becomes very close to zero, which may prevent us from obtaining a solution for $\mathbf{Z_0^+} - \mathbf{Z_0^-}$ (Siewert 2000). DISORT removes this singularity by slightly changing the value of μ_0 when the above condition is present as discussed by Lin et al. (2015).

7.4 Fourier Expansion of the Surface BRDF

The coefficients $\rho_d^m(\mu, -\mu')$ in the Fourier expansion of the surface bidirectional reflectivity in (60) are computed from the defining equation

$$\rho_d^m(\mu, -\mu') = (2 - \delta_{m0}) \frac{1}{\pi} \int_0^\pi \rho_d(\mu, -\mu'; \phi - \phi') \cos m(\phi - \phi') \, \mathrm{d}(\phi - \phi'), \tag{119}$$

where we exploited the assumption that ρ_d is an even function of the azimuth difference $\Delta\phi = \phi - \phi'$ so the integral in (119) is evaluated on the interval $[0, \pi]$ instead of $[-\pi, \pi]$. In DISORT, this integral is evaluated by an N_g-point Gaussian quadrature with weights w_k

$$\rho_d^m(\mu, -\mu') \approx (2 - \delta_{m0}) \frac{1}{\pi} \sum_{k=1}^{N_g} w_k \rho_d(\mu, -\mu'; \Delta\phi_k) \cos(m\Delta\phi_k). \tag{120}$$

ρ_d^m must be obtained for all combinations of incident and reflection angles [(μ_i, μ_j), (μ_i, μ_0), (μ, μ_j), (μ, μ_0)], that is, (120) must be evaluated $N_{\text{total}} = 2NN_g(N^2 + NN_p + N + N_p)$ times, where N is the number of streams per hemisphere, N_p is the number of user-defined upward directions, and N_g is the number of Gaussian quadrature points in (120). The number N_{total} can be very large. For example, when $N = 64$, $N_g = 256$, $N_p = 9$, N_{total} is 155,484,160! Hence, computation of the Fourier expansion coefficients can become very expensive, especially when DISORT is called many times, for example, to calculate intensities at many different wavelengths as would be the case in a line-by-line code.

The integrand in (119) contains the product of a slowly varying function, ρ_d and a highly oscillating cosine function, $\cos(m\Delta\phi_k)$ whose frequency increases with m (and thus the numbers of streams). In order to adequately sample the integrand and to maintain high accuracy, great care should be taken in choosing N_g. On one hand, the quadrature points must be chosen such that ρ_d is sampled adequately when m is small, since the presumed variability of the integrand in (119) is determined by ρ_d. On the other hand, when m is large the variability of the integrand will be due to the highly oscillating component, $\cos(m\Delta\phi_k)$. This situation is illustrated in Fig. 5. The top panel shows $\rho_d(\mu, -\mu, \Delta\phi)$ as a function of $\Delta\phi$ for three values of the direction of reflection μ (0, 0.5, and 0.9) for incoming beam direction $-\mu = 0.5$, as calculated from Hapke's bidirectional reflectance model (Hapke 1993, Eq. 8.89, with parameters $h = 0.06$, $B_0 = 1$ and $w = 0.19$). These functions are not too complex and could be integrated with a low-order quadrature. The bottom panel of Fig. 5 shows the product $\rho_d(\mu, -\mu, \Delta\phi)\cos(m\Delta\phi)$ for three values of m (2, 4, and 16) for the pair of incident/reflection directions ($\mu = -\mu = 0.5$) that exhibits the largest variability with $\Delta\phi$ because it includes the "hot spot". It is clearly seen that the frequency of oscillations increases with increasing m, and therefore the order of quadrature required should also be increasing.

According to the sampling theorem, a cosine function should be sampled at its Nyquist critical frequency; because such sampling guaranties that it will be sampled at its positive peaks and negative troughs (Press et al. 1992). So, N_g should be at least twice as large as the number of streams ($2N$). Although the sampling theory stated in this form applies to evenly spaced samples, and thus the quadrature points satisfying the above requirement strictly apply only to Newton–Cotes quadrature formulas, we may try adopting them for the Gaussian quadrature as well. Combining requirements for small and large m we have $N_g \geq \max(N_{\min}, 4N)$, where $N_{\min}$ is the minimum number of quadrature points needed to adequately

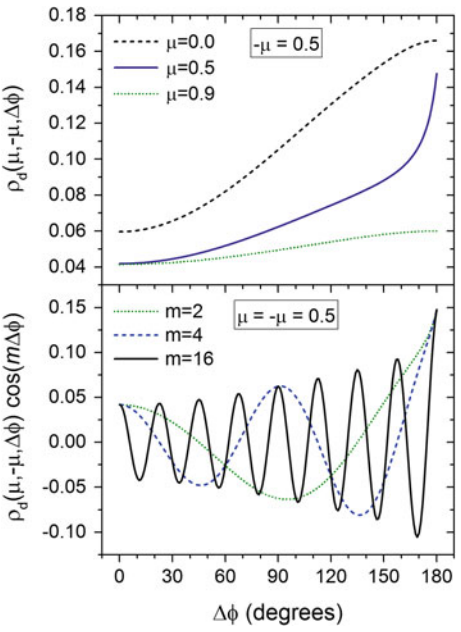


Fig. 5 Illustration of the integrand evaluated for computing the Fourier expansion coefficients of the surface bidirectional distribution function ρ_d. Hapke's BRDF model (Hapke 1993) is used. *Top panel* ρ_d as function of the relative azimuth angle $\Delta\phi$ for three values of reflection directions $\mu = 0.0, 0.5, 0.9$ with incident direction $-\mu = 0.5$. *Bottom panel* product of ρ_d and cos $(m\ \Delta\phi)$ as a function of $\Delta\phi$ for $m = 2, 4, 16$ and for equal incident and reflection directions ($\mu = -\mu = 0.5$). Note the increase of oscillations with increasing m

represent the integrand for small m. In Fig. 6 we show, as an example, the Fourier expansion coefficients $\rho_d^m(\mu, -\mu')$ for $m = 0, \ldots, 127$ (corresponding to 128 streams) obtained using different number of Gaussian quadrature points $N_g = 50$, 100, 200, and 256 in (120). (Hapke's BRDF model with the same parameters as in Fig. 5 is used, except $w = 0.6$.) $\rho_d^m(\mu, -\mu')$ decreases rapidly with increasing m in all cases, however, with $N_g = 50$ it starts widely oscillating after about $m = 25$, indicating that the number of quadrature points used in (120) is not sufficient to approximate the integral in (119). The erroneous oscillation kicks in at larger m value for $N_g = 100$ and 200, but it is only absent when $N_g = 256$, that is when N_g equals twice the number of streams.

We note that the use of Gaussian quadrature in (119) may not be the best choice for evaluating the integral because, as discussed above, the integrand oscillates too much for large m. For such integrands, any conventional quadrature is likely to choose an unrepresentative sample set of values. One obvious method to evaluate the integral is to break up the integrand for intervals where it has the same sign and then apply the Gaussian quadrature separately for each interval. This procedure is, however, likely to be inefficient due to the large number of integrals that should be evaluated and then summed up. A better way to evaluate the integral in (119) may be to use a Filon-type method, which is an extension of the Filon quadrature (Filon 1928). The Filon quadrature approximates the slowly varying component by a polynomial on separate intervals, called panels, then knowing the exact formulas for integrating $x^n \cos(m\ x)$ and summing over all panels obtains a better value for the integral than the standard Gaussian quadrature could give. Unfortunately, the

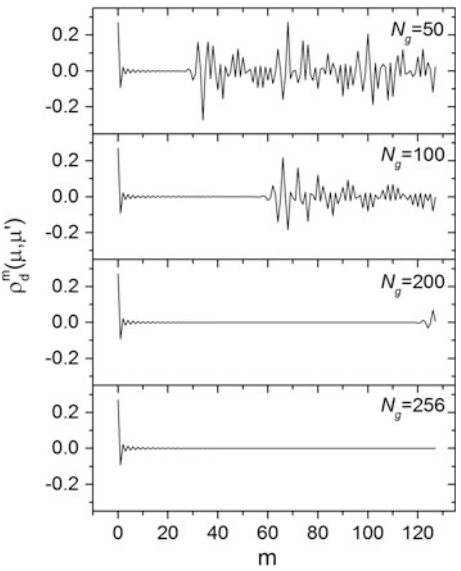


Fig. 6 Fourier expansion coefficients $\rho_d^m(\mu, -\mu')$ of Hapke's BRDF for m values corresponding to 128 streams, that is for $m = 0$-127, obtained using different number of Gaussian quadrature points $N_g = 50$, 100, 200, and 256 in Eq. (120). Note the wild, erroneous oscillation at large m value only disappears when N_g equals twice the number of streams ($N_g = 256$)

quadrature points selected in the Filon-type method are not a good choice for small m, but they can be chosen such that they approach the Legendre points as the frequency of oscillations approaches zero (Iserles and Nørsett 2004). Asheim et al. (2014) discusses how to select the quadrature points in an optimal way. Another choice for evaluating the integral in (119) would be to use Levin-type methods that are further developments of the Filon method, and is based on solving an associated differential equation (Levin 1996). Both of these methods are asymptotic methods. An added benefit of these methods is that, unlike classical Gaussian quadratures, whose error increases with $m \to \infty$, the accuracy of both methods increases with growing oscillation. A unified description of these and other asymptotic methods is presented in Iserles et al. (2006).

7.5 $\omega = 1$ Special Case

The solution (32) presented for the sum of eigenvectors $\mathbf{G}^+ + \mathbf{G}^-$ is only valid when $k \neq 0$. Indeed, $k = 0$ causes apparent infinities in (32). Such a situation happens for $m = 0$ when the single-scattering albedo $\omega = 1$, as we show below for the two-stream case. (For $2N > 2$, cf. the discussion in Chandrasekhar 1960, Chap. 6, p. 212.)

The two-stream approximation is obtained by setting $N = 1$ in (14) and (15), which yields two coupled differential equations ($\mu_{-1} = -\mu_1$, $w_{-1} = w_1 = 1$), which in matrix form is written for the $m = 0$ Fourier component as

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{bmatrix} I^+ \\ I_- \end{bmatrix} = \begin{bmatrix} -\alpha & -\beta \\ \beta & \alpha \end{bmatrix}\begin{bmatrix} I^+ \\ I_- \end{bmatrix} - \begin{bmatrix} Q'^+ \\ Q'^- \end{bmatrix}, \tag{121}$$

where the $m = 0$ superscript has been dropped, and where the following definitions are used

$$\begin{aligned} I^{\pm} &\equiv I(\tau, \pm\mu_1), \\ Q^{\pm} &\equiv Q(\tau, \pm\mu_1), \\ Q'^{\pm} &= \pm Q^{\pm}/\mu_1, \\ \alpha &= [D(\mu_1, \mu_1) - 1]/\mu_1 = \left[\frac{\omega}{2}\left(1 + 3g_1\mu_1^2\right) - 1\right]/\mu_1, \\ \beta &= D(\mu_1, -\mu_1)/\mu_1 = \frac{\omega}{2}\left(1 - 3g_1\mu_1^2\right)/\mu_1. \end{aligned}$$

Seeking the solution to the homogeneous version of (121) ($Q'^{\pm}$=0) in the form

$$I^{\pm} = G^{\pm}\mathrm{e}^{-k\tau}, \quad \text{where} \quad G^{\pm} \equiv G(\pm\mu_1),$$

leads to the following algebraic eigenvalue problem:

$$\begin{bmatrix} \alpha & \beta \\ -\beta & -\alpha \end{bmatrix}\begin{bmatrix} G^+ \\ G^- \end{bmatrix} = k\begin{bmatrix} G^+ \\ G^- \end{bmatrix}.$$

Here k is an eigenvalue and G is an eigenvector. Expanding this into two scalar equations,

$$\begin{aligned} \alpha G^+ + \beta G^- &= k G^+, \\ -\beta G^+ - \alpha G^- &= k G^-, \end{aligned}$$

and adding and subtracting these two equations, we find

$$\begin{aligned} (\alpha - \beta)(G^+ - G^-) &= k(G^+ + G^-), \\ (\alpha + \beta)(G^+ + G^-) &= k(G^+ - G^-). \end{aligned}$$

Substitution of the second of these equations into the first one yields

$$(\alpha - \beta)(\alpha + \beta)(G^+ + G^-) = k^2(G^+ + G^-).$$

Canceling the scalar factor $(G^+ + G^-)$ from both sides, and taking the square root yields the two eigenvalues:

$$k_1 = k, \quad k_{-1} = -k,$$
$$k \equiv \sqrt{\alpha^2 - \beta^2} = \frac{1}{\mu_1}\sqrt{(\omega - 1)\,(3\,\omega\,g_1\,\mu_1^2 - 1)}.$$

Note that in the equation above $\omega = 1$ leads to $k = 0$, which in turn prevents us from obtaining a solution for the eigenvectors from (32). Mathematically, there is no special significance of a zero eigenvalue, and any eigenvalue can be shifted away from zero (Press et al. 1992). The physical equivalence of this would be the existence of radiative transfer equations for the special case of $\omega = 1$. However, even though such transfer equations exist, no solutions specific to this case have been implemented in DISORT. Instead, the zero eigenvalue resulting from $\omega = 1$ is avoided by slightly changing ω, as discussed below.

Conservative scattering ($\omega = 1$) cannot be realized in practice because all materials are absorbing (Bohren and Huffman 1983). Thus, $\omega = 1$ can only be realized as an idealized limit from below ($\omega \rightarrow 1-$) as absorption approaches zero. But this limit involves 0/0 singularities which must be treated by L'Hospital's Rule. This limit leads to entirely different functions than the exponentials found in the general-case formula. Having different functions greatly complicates the interface conditions between layers with $\omega = 1$ and those with $\omega < 1$. DISORT deals with the $\omega = 1$ case by "dithering". Dithering simply means replacing $\omega = 1$ by $\omega = 1 - \varepsilon$ wherever it occurs. (ε is taken to be 10 times machine precision or, in the recommended IEEE double precision mode of operation, about 10^{-13}.) This choice basically gets the computer to do the L'Hospital's Rule computation, but digitally instead of analytically. It allows the $\omega < 1$ solution to be used in every layer, vastly simplifying the interface conditions.

7.6 *Simplified Albedo and Transmissivity Computations*

Frequently one wants to know the flux albedo $a(\mu)$ and transmission $t(\mu)$ of the whole medium for many incident beam angles. The conventional procedure to compute $a(\mu)$ and $t(\mu)$ for a particular angle of incidence μ requires computing the angular distribution of the azimuthally averaged intensity, and then the quadrature of this intensity over angle. This procedure requires the computation of the particular solution (Sect. 4.2) for each layer and every direction of incidence considered. This approach can be quite costly in an inhomogeneous (multilayered) medium. DISORT has a special case, which allows one to get $a(\mu)$ and transmission $t(\mu)$ of the whole medium for many incident beam angles at once. Note that this special case excludes thermal sources and non-Lambertian lower boundary reflectivities!

Using the reciprocity principle, Stamnes (1982b) derived simple expressions for the albedo and transmissivity of a vertically inhomogeneous, plane parallel medium lacking thermal sources:

$$\begin{aligned} a(\mu) &= I^0(0,\mu), & t(\mu) &= I^{0*}(0,\mu), \\ a^*(\mu) &= I^{0*}(\tau_L,-\mu), & t^*(\mu) &= I^0(\tau_L,-\mu), \end{aligned} \tag{122}$$

where the asterisk refers to illumination from below and τ_L is the total optical thickness of the inhomogeneous medium. These equations show that (a) the albedo for a given angle μ of parallel beam incidence equals the azimuthally averaged reflected intensity for isotropic unit intensity incident at the top boundary; and (b) the transmissivity equals the azimuthally averaged transmitted intensity for isotropic unit intensity incident at the bottom boundary.

Equation (122) offers substantial computational advantages. By applying an isotropic boundary condition and using (122), one may solve for all desired angles of incidence simultaneously and entirely avoid the computation of the particular solution.

The numerical implementation of this special case involves two steps (Stamnes 1982b). In step (1), the albedo and transmissivity for the so-called "standard problem" (no reflecting lower boundary) for isotropic illumination of unit intensity from above and below are computed using (122). For illumination from above, the following explicit expressions [cf. (73) and (77)] are used to calculate two of the right-hand sides in (122):

$$\left\{ \begin{matrix} I^0(0,+\mu) \\ I^0(\tau_L,-\mu) \end{matrix} \right\} = \sum_{n=1}^{L} \sum_{j=1}^{N} \left\{ \hat{C}^0_{-jn} \frac{G^0_{-jn}(\pm\mu)}{1 \mp k^0_{jn}\mu} \hat{E}^0_{-jn}(\pm\mu) + \hat{C}^0_{+jn} \frac{G^0_{+jn}(\pm\mu)}{1 \pm k^0_{jn}\mu} \hat{E}^0_{+jn}(\pm\mu) \right\}, \tag{123}$$

where the 0-superscript on $\hat{C}$, $\hat{E}$, k refers to the azimuth-independent case ($m = 0$) and $\hat{E}$ is defined in (74)–76) and (78)–(80). Two equations have been collapsed into one here, because the right-hand sides in (123) are identical except for some sign flips; the two left-hand sides are stacked inside the curly brackets. Equation (123) gives us $a(\mu)$ and $t^*(\mu)$. The same formulas also give us $a^*(\mu)$ and $t(\mu)$. Only the constants of integration $\hat{C}$ are different. Thus, only the final step in the solving procedure, which determines these constants, needs to be repeated for illumination from below. In step (2), the solution from step (1) is modified to account for a Lambert (isotropic) reflecting lower boundary.

8 DISORT Test Cases

DISORT has been distributed together with a set of test cases. The purpose of the test cases is to demonstrate the basic capabilities of the code, exercise all features of DISORT thereby testing all subroutines (logical paths), and provide comparisons

with numerical results obtained by other investigators. They also serve as examples (templates) for appropriately setting up DISORT runs.

There are two groups of test problems: accuracy and consistency groups. The accuracy problems compare the results with answers obtained running DISORT entirely in double precision. The latter results are saved in the DISORT test-case file and read in when the test cases are run. The accuracy problems are not chosen to be comprehensive but rather to test extreme values of the input parameters. Some of the test cases were also chosen to be compared to published results (cf. Van de Hulst 1980; Sweigart 1970; Garcia and Siewert 1985). A complete description of all test cases is beyond the scope of this review. Below we only highlight a few of them.

Test problems 1–5 involve single purely scattering layers with successively more difficult phase functions (isotropic, Rayleigh, Henyey-Greenstein, Haze L, and Cloud C.1 phase functions; the last two are from Garcia and Siewert (1985). Subcases are for varying optical depths and single-scattering albedos.

For example, Case 1a, the simplest case of all, deals with a single isotropically scattering layer over a nonreflecting Lambertian surface, illuminated at the top boundary from the direction $\mu_0 = 0.1$ and $\phi_0 = 0^o$ by a parallel beam with $F_0 = \pi$. The optical thickness of the layer is $\tau_L = 0.03125$, and the single-scattering albedo is $\omega = 0.2$. Computations are to be carried out with $2N = 16$ streams. Both intensities and fluxes are requested; they are to be computed at the top and bottom boundaries. Intensities are requested at six user-specified directions with polar angle cosines $\mu = -1.0, -0.5, -0.1, 0.1, 0.5$, and 1.0, and with azimuth $\phi = 0°$. Writing these with the corresponding DISORT variables (in upper case) for input one has:

NSTR = 16 (number of streams, $2N$)
NTAU = 2 (number of optical depths where results are requested)
UTAU(1) = 0.0 (first user-specified optical depth where results are requested; because we want results at the top and bottom of the layer, and because τ must be in increasing order, we set this to zero for the top boundary)
UTAU(2) = 0.03125 (second user-specified optical depth; since we also want results at the bottom of the layer we set UTAU(2) to the total optical depth of the layer τ_L)
NLYR = 1 (we are dealing with a single layer, so we set the number of computational layers to 1)
DTAUC(1) = UTAU(2) (optical thickness of computational layer; it is set this way because in Case 1a UTAU(2) represents the bottom boundary; alternatively, one could write DTAUC(1) = 0.03125)
SSALB(1) = 0.2 (single-scatter albedo of computational layer; index denotes layer number)
NMOM = NSTR (number of phase function moments not including the zeroth moment. It should be greater than or equal to NSTR in problems with scattering.)
PMOM (0, 1) = 1.0; PMOM(1 to NMOM, 1) = 0.0 (array containing the coefficients in Legendre polynomial expansions of phase functions for computational layers. First index is for moments, second index is for layers.

Subroutine GETMOM supplied with the DISORT test cases is used to set the moments.)
USRTAU = .TRUE. (tells DISORT that radiant quantities are to be returned at user-specified optical depths)
USRANG = .TRUE. (tells DISORT that radiant quantities are to be returned at user-specified polar angles)
NUMU = 6 (number of user-defined polar angles)
UMU(1-6) = [−1.0, −0.5, −0.1, 0.1, 0.5, 1.0] (cosines of output polar angles in increasing order)
NPHI = 1 (number of azimuthal angles at which to return intensities)
PHI(1) = 0.0 (azimuthal output angles in degrees)
IBCND = 0 (tells DISORT that request is for the general case not for the special case in which only albedo and transmissivity of the entire medium versus incident beam angle is requested)
UMU0 = 0.1 (polar angle cosine of incident beam)
PHI0 = 0.0 (azimuth angle of incident beam)
FBEAM = π/UMU0 (incident parallel beam at top boundary; the flux is set to π so that DISORT intensities are equal to Van de Hulst's (1980) reflectivities and transmissivities. DISORT reflected and transmitted fluxes must be divided by the incident flux (π) to compare to the "FLUX" column in Van de Hulst.)
FISOT = 0.0 (intensity of top boundary isotropic illumination)
LAMBER = .TRUE. (instructs DISORT the bottom boundary is an isotropically reflecting surface)
ALBEDO = 0.0 (albedo of isotropically reflecting bottom boundary)
PLANK = .FALSE. (instructs DISORT the problem does not include thermal emission)
ONLYFL = .FALSE. (instructs DISORT to compute intensities in addition to fluxes)

For illustration, Fig. 7 plots the intensities $I(\mu,\phi)$ as a function of polar angle cosines μ with azimuth $\phi = 0°$ in the upward direction at the top and in the downward direction at the bottom. The conditions are those in Case 1 and 1b, except the user polar angles range from 0° to 180° in one-degree increments. The intensities are calculated for single-scattering albedos $\omega = 0.2$ (Case 1a) and $\omega = 1.0$ (Case 1b). The strong "limb brightening" of intensities is evident in both cases.

Test case 5a treats a single cloud layer with the highly anisotropic Cloud C.1 phase functions (Garcia and Siewert 1985). Intensities are calculated using 48 streams. This case is an example of using more moments than number of streams. This is achieved by setting the appropriate DISORT input variables as

$$\text{NSTR} = 48$$
$$\text{NMOM} = 299$$

The test case calls the supplied subroutine GETMOM with option 5 to select the Cloud C.1 phase function and to supply the Legendre moments.

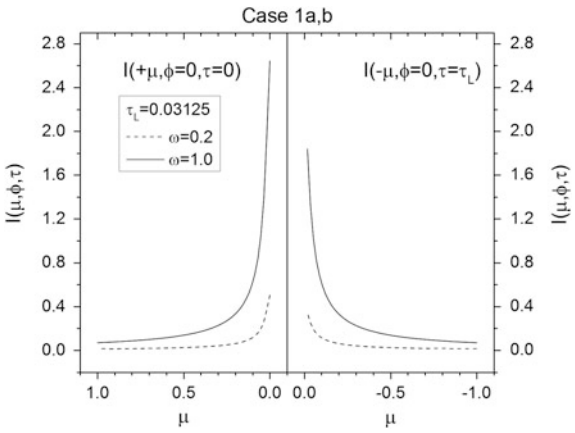


Fig. 7 Intensities $I(\mu, \phi)$ in the upward direction at the top (*left*) and in the downward direction at the bottom (*right*) of a single isotropically scattering layer with total optical thickness $\tau_L = 0.03125$ as a function of polar angle cosines μ with azimuth $\phi = 0°$. Intensities are calculated for single-scattering albedos $\omega = 0.2$ (*dashed line*) and $\omega = 1.0$ (*solid line*). The layer is illuminated at the top by a parallel beam of intensity π from the direction $\mu_0 = 0.1$ and $\phi_0 = 0°$. These results are from Case 1a and 1b of the DISORT test cases, but for every one degree in polar angle (the intensities are given in arbitrary units)

The DISORT results for intensities and fluxes from Case 1 can be compared to those in Table 12 of Van de Hulst (1980). Those from Cases 2–5 can be compared to results in Table 1 of Sweigart (1970), Table 37 of Van de Hulst (1980), and Tables 12–20 of Garcia and Siewert (1985).

As another illustration, Fig. 8 plots the intensities from Case 4c, but for every one degree in polar angle. This case calls for computing intensities at three optical depths $\tau = 0$, 0.5 and 1.0 of a single layer of total optical thickness $\tau_L = 1.0$. The layer scatters radiation according to the Haze-L phase function (Garcia and Siewert 1985) with a single-scattering albedo $\omega = 0.9$, and it is bounded from the bottom by a surface that does not reflect radiation (ALBEDO = 0). The layer is illuminated at the top from the direction $\mu_0 = 0.5$ and $\phi_0 = 0^o$ by a parallel beam with flux $F_0 = \pi$ normal to the beam direction. Figure 8 plots the upward and downward intensities in a polar plot at the top, middle, and bottom of the layer in four azimuth directions $\phi = 0°$, 90°, 180°, and 270°. The plots on the left show the intensity in the principal plane. The maximum downward intensity in the forward-scattering direction ($\mu = -0.5$, $\phi = 0°$) is clearly seen on these plots. The plots on the right show the intensity in the plane perpendicular to the principal plane, and illustrate the symmetry of intensities to the principal plane.

Figure 9 replots the intensity at the middle of the layer as a function of the polar angle starting from the direction straight up, moving to the horizon, and then looking straight down. This plot is to demonstrate the continuity in intensity as direction changes from upward to downward. Figure 9 also plots the intensities at the 32 Gaussian quadrature points (computational polar angle cosines) that resulted

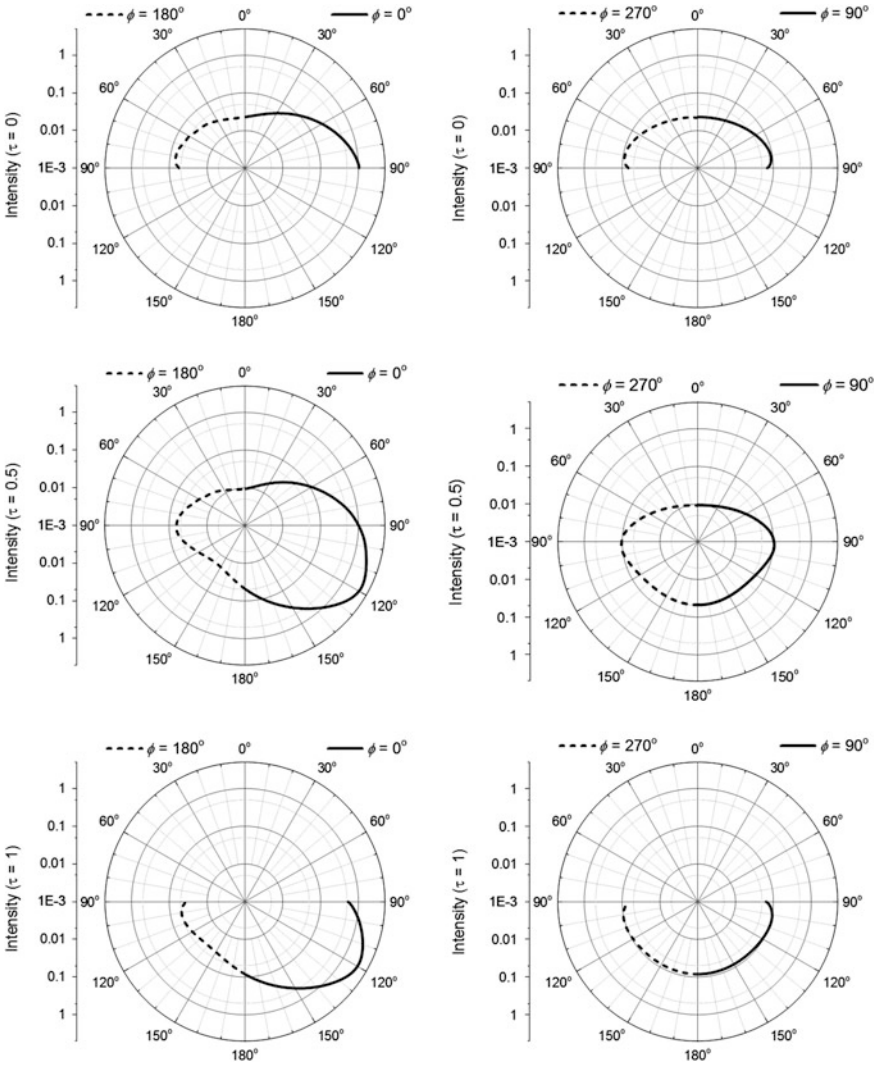


Fig. 8 Intensities computed at three optical depths $\tau = 0.0$, 0.5 and 1.0 of a single layer of total optical thickness $\tau_L = 1.0$. The layer scatters radiation according to the Haze-L phase function with a single-scattering albedo $\omega = 0.9$, and it is bounded from the bottom by a nonreflecting surface. The layer is illuminated at the top from the direction $\mu_0 = 0.5$ and $\phi_0 = 0°$ by a parallel beam with intensity $F_0 = \pi$. Results are from DISORT Test Case 4c, but for every one degree in polar angle

from using 32 streams. The plot demonstrates that the intensities at user-defined polar angles (lines) "fit" well those computed at the quadrature points (symbols); that is, the technique described in Sect. 4.4 for interpolating on the phase function to get intensities at arbitrary directions is implemented correctly (see also Test Case 10 below).

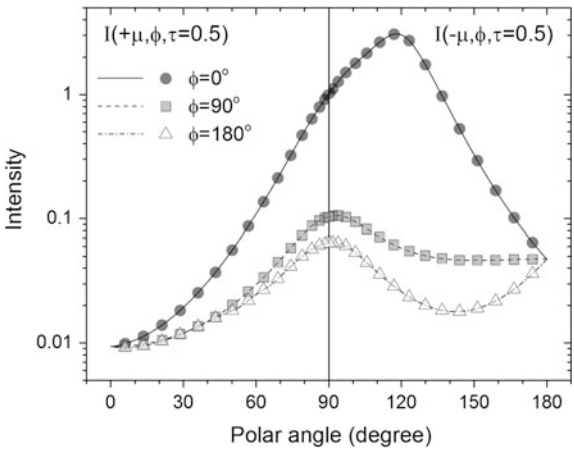


Fig. 9 Intensity as a function of polar angle from the case shown in Fig. 8, but only for the middle of the layer. *Lines* represent intensities at user-defined polar angles (specified at every one degree between 0 and 180°). Symbols denote intensities at the 32 Gaussian quadrature points corresponding to the 32 streams used in the calculation. Note the lines "fit" the symbols well indicating correct calculation of intensity at an arbitrary angle. Also note the continuity in intensity as direction changes from upward to downward

Test problem 6 checks purely absorbing cases with an increasingly complex mix of boundary and internal sources. Subcases include transparent medium over a nonreflecting surface, illuminated with a parallel beam source; the same with less transparent medium; adding Lambertian bottom boundary; replacing the Lambertian bottom boundary with a non-isotropic reflection; adding bottom-boundary emission from a surface; adding diffuse radiation incident at the top boundary; adding internal emission with different temperatures at layer top and at layer bottom; and finally increasing the optical thickness of the layer. The solutions from this test can be compared with exact analytic results. Note that because in accounting for temperature variation across a layer DISORT assumes the Planck function varies linearly in optical depth the temperature gradient should not be large. The current threshold for the gradient is 10 K; when this is exceeded DISORT prints a warning message.

Test problem 7 checks a general emitting/absorbing/scattering medium with one computational layer and an increasingly complex bottom reflectivity condition. Subcases a and b can be compared to results in Tables I and II of Kylling and Stamnes (1992).

Test problems 6 and 7 can be used to set up problems with thermal emissions. For example, test Case 7b shows how to set up calculations of intensities in a narrow spectral interval (2502.99–2703.01 cm^{-1}) from a scattering–absorbing layer with internal thermal emission. The temperature at the top of the layer is 200 K and that at the bottom is 300 K. The layer is assumed to be isolated and have no

radiation incident on it from either boundary. The DISORT input variables corresponding to this situation are:

PLANK = .TRUE. (indicates thermal emission should be included)
BTEMP = 0.0 (temperature of bottom boundary [K])
TTEMP = 0.0 (Temperature of top boundary [K])
TEMPER(0) = 200.0 (temperature at the top of layer)
TEMPER(1) = 300.0 (temperature at the bottom of layer)
WVNMLO = 2702.99 (lower wavenumber of spectral interval [cm^{-1}])
WVNMHI = 2703.01 (upper wavenumber of spectral interval [cm^{-1}])
FBEAM = 0.0 (no parallel beam)
FISOT = 0.0 (no isotropic incident radiation at the top)

Test problem 8 checks an absorbing/isotropic-scattering medium with two computational layers and isotropic incidence at the top boundary. These solutions can also be compared to published results (Table 1 of Ozisik and Shouman 1980).

Test problem 9 checks a general emitting/absorbing/scattering medium with every computational layer different. This case is the most complex test case. Subcases a and b can be compared to results in Tables VI-VII of Devaux et al. (1979). This and test problem 8 can be used as templates to set up multilayer problems. For example, Case 8c does calculations for two layers. The upper layer has optical thickness 1.0 and single single-scatter albedo 0.8. These for the bottom layer are 2.0 and 0.95, respectively. The layers are assumed to scatter isotropically. The corresponding DISORT input variables are:

DTAUC(1) = 1.0
DTAUC(2) = 2.0
SSALB(1) = 0.8
SSALB(2) = 0.95
PMOM (0, 1) = 1.0; PMOM(1 to NMOM, 1) = 0.0
PMOM (0, 2) = 1.0; PMOM(1 to NMOM, 2) = 0.0

Some of the test cases also consider lower boundaries with different complexity, including bidirectional reflection (e.g., cases 6d, 14, and 15). It should be noted that while in DISORT incident (downward) directions have polar angles larger than 90 degrees many BRDF models used in remote sensing define the incident direction less than 90 degrees, that is, the BRDF incident polar angle cosine is positive. The situation is illustrated in Fig. 10. The BRDF and DISORT unit vectors $\bar{\boldsymbol{e}}$ along the incident directions point in opposite directions, and so the relationship between the incident BRDF polar angle θ'_{BRDF} and that of DISORT $\theta'_{\mathrm{DISORT}}$ is $\theta'_{\mathrm{BRDF}} = \pi - \theta'_{\mathrm{DISORT}}$, that is the BRDF incident polar angle is the supplementary of the DISORT polar angle. Similarly, the relationship between the azimuths are $\phi'_{\mathrm{BRDF}} = \phi'_{\mathrm{DISORT}} + \pi$.

Figure 11 shows an example from Case 15a with a bidirectional bottom boundary described by Hapke's model (Hapke 1993). The functional form of this BRDF model is written as (Lin et al. 2015):

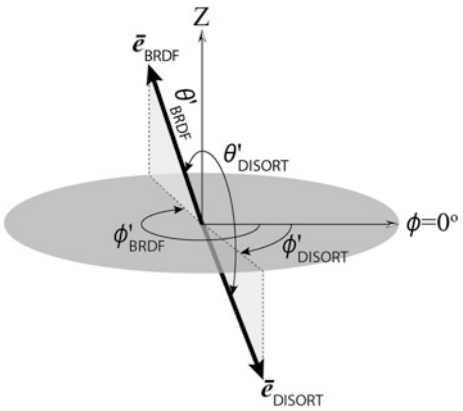


Fig. 10 Polar and azimuth coordinates of incident direction in many popular BRDF models versus those defined in DISORT. BRDF and DISORT unit vectors $\bar{e}$ along the incident directions point opposite directions. $\theta'_{\mathrm{BRDF}} = \pi - \theta'_{\mathrm{DISORT}}$ and $\phi'_{\mathrm{BRDF}} = \phi'_{\mathrm{DISORT}} + \pi$

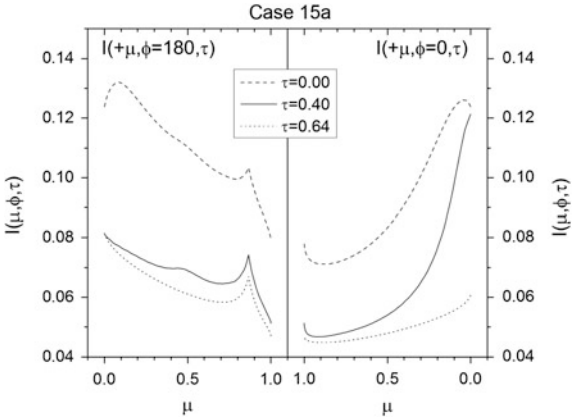


Fig. 11 Upward intensities in the principal plane emerging from a two-layer (aerosol and molecular) atmosphere bounded by a bidirectionally reflecting surface. The molecular layer is on top of the aerosol layer. Intensities are shown at the surface (*dotted line*), just below the top of the aerosol layer (*solid line*), and at the top of the atmosphere (*dashed line*). Results are from DISORT test Case 15a (see text for details)

$$\rho_d(\mu, \mu', \Delta\phi) = \frac{w}{4\pi}\left[\frac{(1+b)p(\cos\alpha) + h_0 h - 1}{\mu + \mu'}\right],$$

where w is the BRDF single-scattering albedo, $\Delta\phi$ is the relative azimuth; α is the phase angle ($\cos\alpha = \mu\mu' - \sqrt{1-\mu^2}\sqrt{1-\mu'^2}\cos(\Delta\phi)$); $p(\cos\alpha)$ is the BRDF phase function (p = 1+0.5 $\cos(\alpha)$); parameters h and h_0 are determined by the so-called albedo factor $\gamma = (1-w)^{1/2}$, and the angles of incidence μ' and reflection μ:

$$h_0 = \frac{1+2\mu'}{1+2\mu'\gamma} \quad \text{and} \quad h = \frac{1+2\mu}{1+2\mu\gamma}.$$

The parameter b accounts for the opposition effect (also called retroreflectance, or "hot spot"); it depends on the angular width parameter h_h of the opposition effect and an empirical factor b_0 that accounts for the finite size of particles:

$$b = \frac{b_0 h_h}{h_h + \tan\left(\frac{\alpha}{2}\right)}.$$

In Case 15a, these parameters are defined as $w = 0.6$, $b_0 = 1.0$, $h_h = 0.06$. In this test case the above bidirectionally reflecting surface bounds from below an atmosphere composed of an aerosol layer and a molecular layer on top of the aerosol layer. Both layers have the same optical thicknesses DTAU(1) = DTAU(2) = 0.32, and both conservative scatterers ω = SSALB(1) = SSALB(2) = 1.0. The Legendre moments PMOM again are from the supplied subroutine GETMOM, and they correspond to the Rayleigh phase function for the top layer and an aerosol phase function for the bottom layer. The microphysical properties of aerosol are those described in Kokhanovsky et al. (2010). The layers are illuminated at the top by a parallel beam of unit flux F_0 = FBEAM = 1.0 normal to the beam from the direction μ_0 = UMU0 = cos(π/6) and azimuth ϕ_0 = PHI0 = 0.0. As an example, Fig. 11 plots the upward intensities in the principal plane at the surface, just below the top of the aerosol layer and at the top of the Rayleigh layer. At all three levels the intensity rapidly decreases in the back-scatter plane ($\phi = 180°$) as the polar angle cosine increases until the hot spot is reached at $\mu = \mu_0$, where it increases and then starts falling again. The intensity is the smallest around the direction $\mu = 1$, and it sharply increases in the forward-scatter plane ($\phi = 0°$) all the way to the horizon ($\mu = 0$). Note how the intensity at the horizon changes at the different levels. At the surface the horizon in the back-scatter direction is "brighter" than in the forward-scatter direction. The aerosol layer changes this behavior by significantly increasing the intensity in the forward-scatter plane above that in the backward-scattering direction. The Rayleigh-scattering layer on top, however, decreases this asymmetry.

The aim of the problems in the consistency group is to compare the results obtained on two successive calls to DISORT, for identical problem structures but taking differing paths through the program. For example, one test problem (Problem 10) checks whether computational polar angles (defined by the Gaussian quadrature for given number of streams) and user-specified polar angles give identical results for the same set of angles. This is done by running DISORT once with the option USRANG = FALSE for a given number of streams, which instructs the code to return intensities at computational polar angles, and once with option USRANG = TRUE, and specifying user angles UMU that correspond to the Gaussian quadrature values for the same number of streams as in the first run. Another problem (Problem 11) checks that identical answers are obtained when user levels in a single layer are replaced by computational levels. In this test, DISORT is first run for a single computational layer of optical thickness 1, and returns intensities and fluxes at four user-specified optical depths (0.00, 0.05, 0.50, 1.00). In the second run, the medium is composed of three layers with single-scattering albedos and phase functions identical in each layer to those in the first run, and optical thicknesses 0.05, 0.45, and 0.5. With this choice of optical thicknesses the cumulative optical depths at the layer boundaries are the same as the

user-specified optical depths in the first run. Answers are then requested at the layer boundaries and compared to those obtained in the first run.

9 A Brief History of DISORT

The first version of DISORT, v1.0, was released in the fall of 1988. It built upon the code developed by Si-Chee Tsay in his thesis (Tsay 1986) under Knut Stamnes. It computed intensity (at either user-specified or computational polar angles and optical depths), mean and azimuthally averaged intensity, direct beam flux, diffuse down- and up-flux, flux divergence, and albedo and transmissivity as a function of incident beam angle. The surface BRDF, similar to the phase function, depended only on the angle between incident and reflected beams, and was expanded into a series of Legendre polynomials. The coefficients were inputs, and as such they were expected to be supplied by the user.

DISORT v1.1, which was considered a finalized version of v1.0, appeared in January 1993. Relative to v1.0 this version included small bug fixes, mostly in ASYMTX; better handling of zero optical depth layers; and other mostly cosmetic changes (printing formats, etc.). Skipping computations for highly absorbing layers was also eliminated when thermal emission was present.

Version 1.2 was released in February 1997 and contained many small cosmetic changes; reorganization of some key subroutines; and improved printing in test problems in order to reduce trivial differences when comparing two outputs. To improve efficiency, calculation of square roots needed in computing the associated Legendre polynomials was moved up to the main DISORT code. Internal code documentation of the subroutine that handles the special case to get only albedo and transmissivity of the entire medium as a function of incident beam angle, and the subroutine that calculates the coefficient matrix was improved.

Version 1.3 was made available to the public in March 2000. In this release, LAPACK replaced LINPACK to do linear algebra, except for the eigenvalue problem. Calculation of machine constants was also changed by calling Fortran-90 intrinsic functions.

While version 1.3 was being developed, independent work was taking place that resulted in version 2.0beta. This version was also released in March 2000, around the same time as v1.3. Version 2.0beta was based on the earlier v1.2, and as such, instead of LAPACK, it still used LINPACK to do linear algebra. Version 2.0beta introduced the Nakajima–Tanaka TMS/IMS intensity corrections (Sect. 6.2) to improve the calculation of intensities. No such corrections had been developed for the azimuthally averaged intensity, so v2.0beta dropped this quantity from the output (we note that although versions prior to V2.0beta did not include the Nakajima-Tanaka intensity correction accurate intensities could still be calculated by turning off the δ–M transformation, provided the increase in computational burden could be tolerated). Version 2.0beta also implemented a more realistic surface BRDF (Sect. 4.5). The latter was necessary because, as pointed out, e.g., by

Godsalve (1995), even in the absence of preferred directions the surface BRDF cannot be assumed to be a function of phase angle only. The distribution of v2.0beta included, as an example, the BRDF developed by Hapke (1993) to model the reflectance of planetary or lunar surfaces.

Calculations of the Fourier expansion coefficients of the surface BRDF were implemented in v2.0beta somewhat inefficiently in that ρ_d was unnecessarily evaluated for each Fourier term. This inefficiency was addressed in version 2.1. In that version, $\rho_d(\mu, -\mu'; \Delta\phi_k)$ in (120) is only evaluated once for $m = 0$, and kept in memory for use with $m > 0$. This change led to an almost fivefold increase in computational speed relative to version 2.0beta. The execution time was further reduced by not calculating the BRDF Fourier coefficients when the surface was the same in repeated runs; it was only evaluated when $m = 0$ or when the number of streams or the surface BRDF was changed.

The number of Gaussian quadrature points N_g used in the computation of Fourier expansion coefficients of the surface BRDF was set to 50 in v2.0beta. Fifty points were sufficient for moderately anisotropic BRDF with low number of streams ($2N \approx 24$), but led to large errors for large $2N$, as discussed in Sect. 7.4. In v2.1 N_g was made a function of the number of streams, and was set to $N_g \geq \max(N_{\min}, 4N)$ with $N_{\min} = 50$.

Version 2.1 also implemented LAPACK to do linear algebra (except for the eigenvalue problem), just like it was done in v1.3. Version 2.1 was provided only to a handful of users for testing; it was not released publicly.

The shortcomings related to the BRDF implementation in v2.0beta mentioned above had been documented in some forms by Laszlo et al., (2010), Stamnes (2011) and Lin et al. (2015), but code with fixes were not publicly released until Version 3.0 in 2015. In v3.0 the Fourier coefficients of the surface BRDF are prepared "offline", thus avoiding the repeated internal computations done in v2.0beta. The "penalty" for the user is that they must compute them. Fortunately, the v3.0 package contains routines to do this for Hapke's model (Hapke 1993) and for three more commonly used BRDFs: Ross–Li (Roujean et al. 1992), Rahman–Pinty–Verstraete (RPV) (Rahman et al. 1993), and Cox–Munk (Cox and Munk 1954). A large enough number of terms in the Fourier expansion of the BRDF is employed to guarantee accurate values of the expansion coefficients; default is 200, but it can be easily changed since it is now an input controllable by the user. These improvements in the treatment of the BRDF have led to a several-fold increase in speed. The intensity correction now also includes the reflection of the direct attenuated beam from the lower boundary (117), which gives a more accurate single-scattering correction. The efficiency for beam sources has been further improved by reducing by a factor of two the dimension of the linear system of equations for the particular solutions (Sect. 4.2), and by replacing the LINPAK routines by LAPACK. In addition, the stability of beam sources has been improved by removing a singularity occurring when the square of the cosine of the incident beam angle is too close to the reciprocal of any of the eigenvalues. These upgrades in v3.0 improved the accuracy, efficiency, and stability of the algorithm.

10 Summary

DISORT implements the discrete ordinate method for solving the transfer of monochromatic unpolarized radiation in non-isothermal, vertically inhomogeneous, but horizontally homogeneous media. The physical processes included are thermal emission, scattering with an arbitrary phase function, absorption, and surface bidirectional reflection. The system may be driven by parallel or isotropic diffuse radiation incident at the top boundary, as well as by internal thermal sources and thermal emission from the boundaries.

The discrete ordinate method implemented in DISORT follows the classical approach of transforming the radiative transfer equation, which is an integro-differential equation, into a system of ordinary differential equations by expanding the phase function into a series of Legendre polynomials and the intensity into a Fourier cosine series. The resulting system of equations is then written in matrix form and the solution of the homogeneous equations is formulated as a standard algebraic eigenvalue problem, which is solved by using a robust linear algebra solver (Stamnes and Dale 1981; Stamnes and Swanson 1981). To make the solution unconditionally stable for arbitrarily large optical depths a scaling transformation is applied (Stamnes and Conklin 1984). Highly anisotropic phase functions are treated by the δ–M method (Wiscombe 1977). Because the δ–M method "truncates" the forward peak of the phase function and alters (scales) the optical depth and Legendre moments, the intensities are less accurate. DISORT corrects the intensities by implementing the Nakajima–Tanaka corrections (Nakajima and Tanaka 1988).

DISORT returns radiances, fluxes, and mean intensities at arbitrary, user-specified angles and levels. The levels need not be subsets of the computational levels necessary to resolve the medium, nor need the angles to be subsets of the quadrature angles necessary to do the integrals over angle. For example, it may require 10 levels to resolve the medium and 16 quadrature angles to do the integrals over angle accurately, but the user can ask for intensities at just two arbitrary levels and one arbitrary angle.

Because the DISORT solutions are closed-form analytic functions of optical thickness computation time is basically independent of optical thickness; but it does depend linearly on the number of computational layers needed to resolve the vertical structure. Since the DISORT solutions are analytic, the number of computational layers can be determined by the actual structure of the problem rather than, as in some other methods, by mere numerical constraints. The only exception to this is that temperature changes across layers should be less than about 10 K to avoid significant errors in the thermal emission computation. In general the number of computational layers will be determined by how accurately one wants to resolve the optical properties (absorption/scattering characteristics) of the medium, which are taken to be constant across any layer.

DISORT has become popular in the atmospheric science and other communities since its introduction in 1988. It has been implemented as the radiative transfer

solver in larger models, e.g., MODTRAN (Berk et al. 1998), SBDART (Ricchiazzi et al. 1998), Streamer (Key and Schweiger 1998). It has been applied to terrestrial (e.g., Tsay et al. 1989, 1990; Tsay and Stamnes 1992) and other planetary (Lindner, 1988) atmospheric radiative transfer problems. It has also been used with other techniques, like the perturbation theory to solve cloud related independent-pixel solar radiative transfer computations which have been embedded in and built around DISORT (Jerg and Trautmann 2007). It has also become a kind of standard against which to compare other models. For example, DISORT has been combined with the Line-by-Line Radiative Transfer Model (LBLRTM) (Clough et al. 2005) to evaluate the Community Radiative Transfer Model (CRTM) developed by the U.S. Joint Center for Satellite Data Assimilation (Ding et al. 2011). DISORT, which is written in Fortran, has also been converted into the C programming language with dynamic memory allocation and consistent use of double precision throughout the code (Buras et al. 2011).

DISORT has been designed to be an example of good scientific software with clear in-code documentation, as well as to be a scientific tool. A substantial effort has been made to make it numerically well-conditioned, error-resistant, user-friendly, and to take advantage of robust software tools. The DISORT distribution includes an extensive test suite designed to verify the program against published results, and for internal consistency. The careful attention to software design has been just as important in DISORT's popularity as its algorithmic content.

Early versions (up to v2.0beta) of the DISORT Fortran source code are available at ftp://climate1.gsfc.nasa.gov/pub/wiscombe/Multiple_Scatt/. The more recent version, v3.0, can be downloaded from http://lllab.phy.stevens.edu/disort/. DISORT v3.0 (Lin et al. 2015) provides important upgrades that improve the accuracy, efficiency, and stability of the algorithm. Compared with v2.0beta (released in 2000) these upgrades include (i) a redesigned BRDF computation that improves both speed and accuracy, (ii) an improved treatment of the single-scattering correction that takes boundary reflection into account, and (iii) additional efficiency and stability upgrades for beam sources. Finally, a correction for Earth curvature effects has been included in DISORT3 by implementing a pseudospherical treatment.

Acknowledgments The authors thank Z. Lin, S. Stamnes, L. Rokke, M. Zhou and H. Liu for their critical reading of and useful comments on an earlier version of the manuscript, and A. Kokhanovsky for inviting us to write this review and for his patience. IL acknowledges the assistance of K. Laszlo with typing in many of the equations.

Disclaimer

The contents of this paper are solely the opinions of the authors and do not constitute a statement of policy, decision, or position on behalf of the U.S. National Oceanic and Atmospheric Administration (NOAA) or the U.S. Government.

References

Anderson E, Bai Z, Bischof C, Demmel I, Dongarra J, Du Croz J, Greenbaum A, Hammarling S, McKenney A, Ostrouchov S, Sorensen D (1999) LAPACK user's guide, 3rd ed. Society for Industrial and Applied Mathematics, 3600 University City Science Center, Philadelphia, Pennsylvania, ISBN: 978-0-89871-447-0, 19104-2688

Asano S (1975) On the discrete ordinates method for radiative transfer. J. Meteor Soc 53:92–95

Asheim A, Deano A, Huybrechs D, Wang HY (2014) A gaussian quadrature rule for oscillatory integrals on a bounded interval. Discret Contin Dyn Syst 34(3):883–901

Berk A, Bernstein LS, Anderson GP, Acharya PK, Robertson DC, Chetwynd JH, Adler-Golden SM (1998) MODTRAN cloud and multiple scattering upgrades with application to AVIRIS. Rem Sens Envir 65(3):367–375. doi:10.1016/S0034-4257(98)00045-5

Bohren C, Huffman D (1983) Absorption and scattering of light by small particles. Wiley, New York

Buras R, Dowling T, Emde C (2011) New secondary-scattering correction in DISORT with increased efficiency for forward scattering. J Quant Spectrosc Radiat Transfer 112(12):2028–2034. doi:10.1016/j.jqsrt.2011.03.019

Chandrasekhar S (1960) Radiative transfer. Dover Publications, New York 393 p

Clough SA, Shephard MW, Mlawer EJ, Delamere JS, Iacono MJ, Cady-Pereira K, Boukabara S, Brown PD (2005) Atmospheric radiative transfer modeling: a summary of the AER codes. J Quant Spectrosc Radiat Transfer 91(2):233–244. doi:10.1016/j.jqsrt.2004.05.058

Cowell WR (ed) (1980) Sources and developments of mathematical sofware. Prentice Hall, Englewood Cliffs

Cox C, Munk W (1954) Measurement of roughness of the sea surface from photographs of the sun's glitter. J Opt Soc Am 44(11):838–850

Devaux C, Grandjean P, Ishiguro Y, Siewert CE (1979) On multi-region problems in radiative transfer. Astrophys Space Sci 62:225–233

Ding S, Yang P, Weng F, Liu Q, Han Y, van Delst P, Li J, Baum B (2011) Validation of the community radiative transfer model. J Quant Spectrosc Radiat Transfer 112(6):1050–1064. doi:10.1016/j.jqsrt.2010.11.009

Dongarra J, Moler C, Bunch J, Stewart GW (1979) LINPACK User's guide. Society for Industrial and Applied Mathematics (SIAM) Press, Philadelphia, pp xx+344. ISBN 978-0-89871-172-1

Filon LNG (1928) On a quadrature formula for trigonometric integrals. Proc Roy Soc Edinburgh Sect A 49:38–47

Garcia RDM, Siewert CE (1985) Benchmark results in radiative transfer. Transp Theory Stat Phys 14:437–483

Godsalve C (1995) The Inclusion of Reflectances with Preferred Directions in Radiative-Transfer Calculations. J Quant Spectrosc Radiat Transfer 53(3):289–305

Godsalve C (1996) Accelerating the discrete ordinates method for the solution of the radiative transfer equation for planetary atmospheres. J Quant Spectrosc Radiat Transfer 56(4):609–616. doi:10.1016/0022-4073(96)00075-1

Hapke B (1993) Theory of reflectance and emittance spectroscopy. Cambridge University Press, Cambridge, p 455

Iserles A, Nørsett SP (2004) On quadrature methods for highly oscillatory integrals and their implementation. BIT 44(4):755–772

Iserles A, Nørsett SP, Olver S (2006) Highly oscillatory quadrature: the story so far. In: de Castro A, Gómez D, Quintela P, Salgado P (eds) Numerical mathematics and advanced applications. Springer, Berlin, pp 97–118

Jerg M, Trautmann T (2007) One-dimensional solar radiative transfer: Perturbation approach and its application to independent-pixel calculations for realistic cloud fields. J Quant Spectrosc Radiat Transfer 105(1):32–54. doi:10.1016/j.jqsrt.2006.09.014

Key J, Schweiger AJ (1998) Tools for atmospheric radiative transfer: Streamer and FluxNet. Comput Geosci 24(5):443–451

Kokhanovsky AA et al (2010) Benchmark results in vector atmospheric radiative transfer. J Quant Spectrosc Radiat Transfer 111(12–13):1931–1946

Kylling A, Stamnes K (1992) Efficient yet accurate solution of the linear transport equation in the presence of internal sources: the exponential-linear-in-depth approximation. J Comp Phys 102:265–276

Laszlo I, Stamnes K, Wiscombe WJ, Tsay S-C (2010) Towards generalized boundary conditions in DISORT. In: Extended abstracts of the AMS 13th conference on atmospheric radiation, 28 June–2 July 2010, Portland, Oregon. https://ams.confex.com/ams/13CldPhy13AtRad/techprogram/paper_171283.htm

Levin D (1996) Fast integration of rapidly oscillatory functions. J Comput Appl Math 67(1): 95–101

Lin Z, Stamnes S, Jin Z, Laszlo I, Tsay SC, Wiscombe WJ, Stamnes K (2015) Improved discrete ordinate solutions in the presence of an anisotropically reflecting lower boundary: upgrades of the DISORT computational tool. J Quant Spectrosc Radiat Transfer 157:119–134. doi:10.1016/j.jqsrt.2015.02.014

Lindner B (1988) Ozone on mars: the effects of clouds and airborne dust. Planet Space Sci 36:125–144

Liou K-N (1973) A numerical experiment on Chandrasekhar's discrete-ordinate method for radiative transfer: applications to cloudy and hazy atmospheres. J Atmos Sci 30(7):1303–1326

Nakajima T (2010) A retrospective view on "algorithms for radiative intensity calculations in moderately thick atmospheres using a truncation approximation" by Teruyuki Nakajima and Masayuki Tanaka (1988). J Quant Spectrosc Radiat Transfer 111(11):1651–1652. doi:http://dx.doi.org/10.1016/j.jqsrt.2010.01.002

Nakajima T, Tanaka M (1986) Matrix formulations for the transfer of solar radiation in a plane-parallel atmosphere. J Quant Spectrosc Radiat Transfer 35:13–21

Nakajima T, Tanaka M (1988) Algorithms for radiative intensity calculations in moderately thick atmospheres using a truncation approximation. J Quant Spectrosc Radiat Transfer 40:51–69

Ozisik M, Shouman S (1980) Source function expansion method for radiative transfer in a two-layer slab. J Quant Spectrosc Radiat Transfer 24:441–449

Parlett and Reinsch (1969) Balancing a matrix for calculation of eigenvalues and eigenvectors. Num Math 13:293–304

Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) Numerical recipes in Fortran: the art of scientific computing, 2nd edn. Cambridge University Press, Cambridge 963 p

Rahman H, Pinty B, Verstraete MM (1993) Coupled surface-atmosphere reflectance (CSAR) model 2. Semiempirical surface model usable with NOAA advanced very high resolution radiometer data. J Geophys Res 98(D11):20791–20801

Ricchiazzi P, Yang S, Gautier C, Sowle D (1998) SBDART: a research and teaching software tool for plane-parallel radiative transfer in the earth's atmosphere. Bull Am Meteorol Soc 79 (10):2101–2114. doi:10.1175/1520-0477(1998)079<2101:SARATS>2.0.CO;2

Roujean J-L, Leroy M, Deschamps P-Y (1992) A bidirectional reflectance model of the Earth's surface for the correction of remote sensing data. J Geophys Res 97(D18):20455–20468

Siewert CE (2000) A concise and accurate solution to Chandrasekhar's basic problem in radiative transfer. J Quant Spectrosc Radiat Transfer 64(2):109–130. doi:10.1016/S0022-4073(98)00144-7

Stamnes K (1982a) On the computation of angular distributions of radiation in planetary atmospheres. J Quant Spectrosc Radiat Transfer 28:47–51

Stamnes K (1982b) Reflection and transmission by a vertically inhomogeneous planetary atmosphere. Planet Space Sci 30:727–732

Stamnes K (1986) The theory of multiple scattering of radiation in plane parallel atmospheres. Rev Geophys 24:299–310

Stamnes K, Conklin P (1984) A new multi-layer discrete ordinate approach to radiative transfer in vertically inhomogeneous atmospheres. J Quant Spectrosc Radiat Transfer 31:273–282

Stamnes K, Dale H (1981) A new look at the discrete ordinate method for radiative transfer calculation in anisotropically scattering atmospheres. II: Intensity computations. J Atmos Sci 38:2696–2706

Stamnes K, Swanson RA (1981) A New look at the discrete ordinate method for radiative transfer calculations in anisotropically scattering atmospheres. J Atmos Sci 38(2):387–399. doi:10.1175/1520-0469(1981)038<0387:ANLATD>2.0.CO;2

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

Stamnes K, Tsay S-C, Nakajima T (1988b) Computation of eigenvalues and eigenvectors for discrete ordinate and matrix operator method radiative transfer. J Quant Spectrosc Radiat Transfer 39:415–419

Stamnes K, Tsay S-C, Wiscombe WJ, Laszlo I (2000) DISORT, a general-purpose Fortran program for discrete-ordinate-method radiative transfer in scattering and emitting layered media: documentation of methodology. NASA Technical Report, version 1.1

Stamnes S (2011) Calculations of the bidirectional reflectance distribution function (BRDF) of a planetary surface. MS thesis, Department of Physics and Engineering Physics, Stevens Institute of Technology, Hoboken, New Jersey, USA

Sweigart A (1970) Radiative transfer in atmospheres scattering according to the Rayleigh phase function with absorption. Astrophys J 22:1–80

Sykes JB (1951) Approximate integration of the equation of transfer. Mon Not R Astron Soc 111 (4):377–386

Thomas GE, Stamnes K (1999) Radiative transfer in the atmosphere and ocean. Cambridge University Press, Cambridge

Tsay S-C (1986) Numerical study of the atmospheric radiative transfer process with application to the Arctic energy balance. Ph.D. thesis, Alaska University, Fairbanks

Tsay S-C, Stamnes K, Jayaweera K (1989) Radiative energy budget in the cloudy and hazy Arctic. J Atmos Sci 46:1002–1018

Tsay S-C, Stamnes K, Jayaweera K (1990) Radiative transfer in stratified atmospheres: development and verification of a unified model. J Quant Spectrosc Radiat Transfer 43:133–148

Tsay S-C, Stamnes K (1992) Ultraviolet radiation in the Arctic: the impact of potential ozone depletions and cloud effects. J Geophy Res 97:7829–7840

Van de Hulst HC (1980) Multiple light scattering, tables, formulas and applications, vol 1 and 2. Academic Press, New York

Wilkinson J (1965) The algebraic eigenvalue problem. Clarendon Press, Oxford

Wiscombe WJ (1976) Extension of the doubling method to inhomogeneous sources. J Quant Spectrosc Radiat Transfer 16:477–489

Wiscombe WJ (1977) The delta–M method: rapid yet accurate radiative flux calculations for strongly asymmetric phase functions. J Atmos Sci 34(9):1408–1422

Community Radiative Transfer Model for Air Quality Studies

Quanhua Liu and Cheng-Hsuan Lu

1 Introduction

Community Radiative Transfer Model (CRTM) (Han et al. 2006), developed at the Joint Center for Satellite Data Assimilation, has being operationally supporting satellite radiance assimilation for weather forecasting and Earth observation space programs. The CRTM has been supporting the Geostationary satellite (GOES)—R and Joint Polar Satellite System (JPSS) Suomi NPP missions for instrument calibration, validation, monitoring long-term trending, and satellite products using a retrieval approach (Liu and Boukabara 2014). At the both National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Prediction (NCEP) and the NOAA National Environmental Satellite, Data, and Information Service (NESDIS) satellite radiance monitoring systems, the CRTM model is applied to simulate satellite observations. The CRTM has been used to assimilate Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol optical depth for air quality forecasting (Liu et al. 2011). Similar to weather forecasting, the air quality forecasting is governed by air fluid dynamic equations based on aerosol sources and the emission of ozone, carbon monoxide, ammonia, and other gases as well as their chemical reactions. Suspended particulate matter

Q. Liu (✉)
Center for Satellite Applications and Research, National Oceanic and Atmospheric Administration, 5830 University Research Court, College Park, MD 20740, USA
e-mail: Quanhua.Liu@noaa.gov

C.-H. Lu
Atmospheric Sciences Research Center, State University of New York, 1400 Washington Ave, Albany, NY 12222, USA

© Springer-Verlag Berlin Heidelberg 2016

A. Kokhanovsky (ed.), *Light Scattering Reviews, Volume 11*,
Springer Praxis Books, DOI 10.1007/978-3-662-49538-4_2

(PM) and ozone at ground level near surface are the critical parameters to issue warning or alert to the public. Ground level or "bad" ozone is not emitted directly into the air, but is created by chemical reactions between oxides of nitrogen (NO_x) and volatile organic compounds (VOC) in the presence of sunlight (Seinfeld and Pandis 2006). The ozone can also be formed during biomass burning (Alvarado et al. 2009). Emissions from industrial facilities and electric utilities, motor vehicle exhaust, gasoline vapors, and chemical solvents are some of the major sources of NO_x and VOC. Breathing ozone can trigger a variety of health problems, particularly for children, the elderly, and people of all ages who have lung diseases such as asthma. Ground-level ozone can also have harmful effects on sensitive vegetation and ecosystems.

Aerosol sources and gaseous emission are the most important information to the air transport model. The NOAA National Air Quality Forecasting Capability (NAQFC), developed by the NOAA Air Resources Laboratory (ARL) and operated by the National Weather Service (NWS), disseminates NOAA's real-time model forecasts and satellite observations of air quality to state and local air quality and public health agencies, as well as the general public. The NWS uses the NAQFS to provide air quality warnings and alerts for large cities in the United States. Air quality and public health managers use these forecasts to inform short-term management decisions and longer term policies to reduce the adverse effects of air pollution on human health and their associated economic costs. The NAQFC operational O_3 and developmental PM2.5 forecasts consist of three model components: the ARL Emission Forecasting System (EFS), the NWS North American Mesoscale (NAM) regional nonhydrostatic meteorological model, and the modified CMAQ model. They all use 12 km horizontal grid spacing and are coupled in a sequential and offline manner with hourly data fed from NAM to Community Multi-scale Air Quality (CMAQ) hourly. While the "operational" O_3 forecasts are guaranteed to be available and be disseminated on time to the general public both graphically and in formatted files following World Meteorological Organization (WMO) standards, the dissemination of the "developmental" PM2.5 forecasts are currently restricted to a selected group of local and state air quality forecasters, the so-called Focus Group Forecasters (FGF) of the NAQFC. The NAQFC is one of the major gateways to disseminate NOAA model prediction and satellite observations of air quality to the public (http://airquality.weather.gov). The accuracy of NAQFC forecasts depend on EFS, NAM, and CMAQ in about the same magnitude of forcing. NAM is quite advanced in utilizing satellite data to nudge weather forecast through its data assimilation schemes (Lee and Liu 2014). However, until recently NAQFC has not assimilated the atmospheric composition data from satellite retrievals to improve its emission, transport, and chemical transformation estimates.

Aerosols can affect the energy balance of Earth's atmosphere through the absorption and scattering of solar and thermal radiation. Aerosols also affect Earth's climate through their effects on cloud microphysics, reflectance, and precipitation. In

addition, aerosols can be viewed as air pollutants because of their adverse health effects. In NOAA NCEP global forecast system (GFS), the representations of aerosols need to better account for these effects. The aerosol distributions in the forecast model are currently prescribed based on a global climatological aerosol database (Hess et al. 1998) and only the aerosol direct effect is considered. The current data assimilation system assumes climatological aerosol conditions. For atmospheric conditions with anomalous high aerosol loading, bias correction and quality control procedures can be compromised due to the unaccounted effects of aerosol attenuation.

An online aerosol modeling capability in Goddard Chemistry Aerosol Radiation and Transport (GOCART) has been developed and implemented within GMAO's GEOS-5 Earth system model (Colarco et al. 2010) and was later coupled with NCEP's NEMS version of GFS to establish the first interactive atmospheric aerosol forecasting system at NCEP, NGAC (Lu et al. 2013). While the ultimate goal at NCEP is a full-up Earth system with the inclusion of aerosol-radiation feedback and aerosol–cloud interaction, the current operational configuration (as in September 2014) is to maintain a dual-resolution configuration with low-resolution (T126 L64) forecast-only system for aerosol prediction (using NGAC) and a high-resolution (T574 L64) forecasting and analysis system for medium-range weather prediction (using GFS). Aerosol fields in NGAC initial conditions are taken from prior NGAC forecast and the corresponding meteorological fields are downscaled from GFS analysis.

The current dual-resolution system mentioned above feeds meteorological fields from GFS to NGAC. In our study, we will upgrade the infrastructure in the dual-resolution system, allowing NGAC aerosol fields fed into GFS. The linkage between NGAC and GFS to be considered include: using NGAC aerosols fields instead of the climatology (Hess et al. 1998) to determine aerosol optical properties in GFS radiation, enabling the GFS data assimilation system to consider NGAC aerosol fields instead of background aerosol loading for satellite radiances calculations, and incorporating NGAC aerosols fields in sea surface temperature (SST) analysis. Standard forecast verification system will be used to determine whether the improved treatment of aerosols will lead to improvement in weather forecasts. Data rejection will be examined to assess whether improved aerosol treatment will lead to better use of satellite data. Diagnosis of energy budget will be conducted to evaluate whether more realistic temporal and spatial representation of aerosols will result in improved energy balance.

Global ammonia (NH_3) emissions have been increasing due to the dramatically increased agricultural livestock numbers together with the increasing use of nitrogen fertilization (Sutton et al. 2014). Atmospheric ammonia has impacts on local scales, acidification and hypertrophication of the ecosystems, and international scales through formation of fine ammonium-containing aerosols. These ammonium aerosols affect Earth's radiative balance and public health. Measurements with daily and large global coverage are challenging and have been lacking partly because the lifetime of NH_3 is relatively short (hours to a day) and partly because it requires high sensitivity for the retrievals. The Cross-Track Infrared Sounder (CrIS) and Infrared Atmospheric Sounding Interferometer (IASI) hyperspectral measurements are good for monitoring NH_3 emissions due to their large daily global coverage and

afternoon overpasses that provide higher thermal contrasts, and hence, higher measurement sensitivities.

NH_3 contributes to the formation of PM2.5 by reaction with nitric and sulfuric acids (HNO_3 and H_2SO_4), which are in turn formed by the photochemical oxidation of nitrogen oxides (NO_x = NO + NO_2) and sulfur dioxide (SO_2), respectively. In the United States, it is estimated that livestock accounts for about 74 % of total anthropogenic NH_3 emissions, with an additional 16 % due to the manufacturing and use of fertilizer (Potter et al. 2010), although these estimates are uncertain. Reductions of NH_3 emissions have been proposed as a cost effective way to improve air quality.

As emphasized above, aerosol sources and gaseous emission are very critical information affecting air quality analysis and prediction. Observations from satellites provide information about global aerosol sources and gaseous emission. However, the sensors onboard satellites measure radiation rather than aerosols or trace gases. To derive aerosol and trace gas from the satellite measurements, inversion technique in retrieval algorithm or radiance assimilation are needed where radiative transfer models play an important role to interpret the meaning of the radiation to geophysical parameters. The CRTM is such an operational radiative transfer model for passive microwave, infrared, and visible sensors.

2 What Are the Common Air Pollutants?

Pollutants are harmful to public health and the environment. The US Environment Protection Agency (EPA) identified six common air pollutants (http://www.epa.gov/air/criteria.html). They are particle (aerosol) pollution (often referred to as particulate matter in the atmosphere), ground-level ozone, carbon monoxide, sulfur oxides, nitrogen oxides, and lead. These pollutants can harm public health and the environment, and cause property damage. Of the six pollutants, particle pollution and ground-level ozone are the most widespread health threats. The first real-time product of NAQFC is ozone. Recently, NAQFC starts to provide PM2.5 (i.e., particulate matter size less than 2.5 μm in diameter) forecasts to early adapter users ("the focus group"), predominantly state environmental agencies that use NAQFC daily products for numeric air quality guidance. EPA calls these pollutants "criteria" air pollutants because it regulates them by developing human health-based and/or environmentally based criteria (science-based guidelines) for setting permissible levels. The set of limits based on human health is called primary standards. Another set of limits intended to prevent environmental and property damage is called secondary standards.

EPA has set National Ambient Air Quality Standards for six principal pollutants (http://www.epa.gov/air/criteria.html), which are called "criteria" pollutants. They are listed below. Units of measure for the standards are parts per million (ppm) by volume, parts per billion (ppb) by volume, and micrograms per cubic meter of air (μg m^{-3}) (see Table 1).

Table 1 Six principal pollutants identified by US Environmental Protection Agency

Pollutant [final rule cite]		Primary/secondary	Averaging Time	Level	Form
Carbon monoxide [76 FR 54294, Aug 31, 2011]		Primary	8-h	9 ppm	Not to be exceeded more than once per year
			1-h	35 ppm	
Lead [73 FR 66964, Nov 12, 2008]		Primary and secondary	Rolling 3 month average	0.15 $\mu g/m^3$ [a]	Not to be exceeded
Nitrogen dioxide [75 FR 6474. Feb 9. 2010] [61 FR 52852, Oct 8, 1996]		Primary	1-h	100 ppb	98th percentile of 1-h daily maximum concentrations, averaged over 3 years
		Primary and secondary	Annual	53 ppb[b]	Annual mean
Ozone [73 FR 16436, Mar 27, 2006]		Primary and secondary	8-h	0.075 ppm[c]	Annual fourth highest daily maximum 8-h concentration, averaged over 3 years
Particle pollution Dec 14, 2012	$PM_{2.5}$	Primary	Annual	12 $\mu g/m^3$	Annual mean, averaged over 3 years
		Secondary	Annual	15 $\mu g/m^3$	Annual mean, averaged over 3 years
		Primary and secondary	24-h	35 $\mu g/m^3$	98th percentile, averaged over 3 years
	PM_{10}	Primary and secondary	24-h	150 $\mu g/m^3$	Not to be exceeded more than once per year on average over 3 years
Sulfur dioxide [75 FR 35520, Jun 22, 2010] [38 FR 25678, Sept 14, 1973]		Primary	1-h	75 ppb[d]	99th percentile of 1-h daily maximum concentrations, averaged over 3 years
		Secondary	3-h	0.5 ppm	Not to be exceeded more than once per year

Detailed description of the following tables is referred to the EPA website (http://www.epa.gov/ttn/naaqs/criteria.html)

[a]Previous standards; current (2008) standards have not been submitted and approved

[b]The level of the annual NO2 standard

[c]The previous (2008) O3 standards additionally remain in effect in some areas

[d]The previous SO2 standards (0.14 ppm 24-hour and 0.03 ppm annual) will additionally remain in effect in certain areas

2.1 Air Pollutants

2.1.1 Ozone

Ozone in the stratosphere is good to protect human from harmful ultraviolet radiation. Oxygen, nitrogen, and ozone block more than 95 % UV radiation from Sun to the Earth's surface. UV light may be divided into three bands: UVC [100–280 nm], UVB (280–315 nm), and UVA (315–400 nm). UVC is the most dangerous, but it is mostly absorbed by oxygen and ozone molecules in the stratosphere and does not reach the Earth's surface. The sun's UV radiation is both a major cause of skin cancer and the best natural source of vitamin D. Therefore, we need a proper UV exposure. Clouds are often opaque that blocks sun's radiation. But, broken clouds may scatter more sun's radiation including UV component due to the three-dimensional radiative transfer. Ozone in the stratosphere is good. The ground-level ozone is bad for public health. Ground-level ozone is not emitted directly into the air, but is created by chemical reactions between oxides of nitrogen (NO_x) and volatile organic compounds (VOC) in the presence of sunlight. Emissions from industrial facilities and electric utilities, motor vehicle exhaust, gasoline vapors, and chemical solvents are some of the major sources of NO_x and VOC. Breathing ozone can trigger a variety of health problems, particularly for children, the elderly, and people of all ages who have lung diseases such as asthma. Ground-level ozone can also have harmful effects on sensitive vegetation and ecosystems.

2.1.2 Particulate Matter

Particulate matter or suspended matter in the atmosphere also known as particle pollution or PM, is a complex mixture of small particles and liquid droplets. Particle pollution is made up of a number of components, including nitrates, sulfates, ammonium, organic chemicals, metals, and soil or dust particles (http://www.epa.gov/airquality/particlepollution/). In the cities, fine particles primarily come from car, truck, bus, and off-road vehicle exhausts, other operations involve biomass burning, heating oil or coal and natural sources. Fine particles also form from the reaction of gases or droplets in the atmosphere from sources such as power plants. The size of particles is directly linked to their potential for causing health problems. EPA is concerned about particles that are 10 μm in diameter or smaller because those are the particles that generally pass through the throat and nose and enter the lungs. Once inhaled, these particles can affect the heart and lungs and cause serious health effects. EPA groups particle pollution into two categories:

- "Inhalable coarse particles," such as those found near roadways and dusty industries, are larger than 2.5 μm and smaller than 10 μm in diameter. Particles less than 10 μm in diameter (PM_{10}) pose a health concern because they can be inhaled into and accumulate in the respiratory system.

- "Fine particles," such as those found in smoke and haze, are 2.5 μm in diameter and smaller. These particles can be directly emitted from sources such as forest fires, or they can form when gases emitted from power plants, industries, and automobiles react in the air. Particles less than 2.5 μm in diameter ($PM_{2.5}$) are referred to as "fine" particles and are believed to pose the greatest health risks.

The EPA categorized air quality to six classes based on PM concentrations http://airnow.gov/index.cfm?action=aqibasics.aqi): 0–50 μg m^{-3} (excellent), 50–100 μg m^{-3} (good), 100–150 μg m^{-3} (lightly polluted), 150–200 μg m^{-3} (moderately polluted), 200–300 μg m^{-3} (heavily polluted), 300–500 (severely polluted), and >500 μg m^{-3} (hazardous). Weather conditions and topography can play an important role in local air quality. Outdoor $PM_{2.5}$ levels are most likely to be elevated on days with little or no wind or air mixing. $PM_{2.5}$ can also be produced by common indoor activities. Some indoor sources of fine particles are tobacco smoke, cooking (e.g., frying, sautéing, and broiling), burning candles or oil lamps, and operating fireplaces and fuel-burning space heaters (e.g., kerosene heaters). One may acquire PM2.5 at nearly real-time online, for example, the air quality in Beijing may be obtained from the website http://aqicn.org/city/beijing/.

2.1.3 Carbon Monoxide

Carbon monoxide (CO) is a colorless, odorless gas emitted from incomplete combustion processes. In urban areas, the majority of CO emissions to ambient air come from mobile sources. In rural areas, biomass burning and forest fires release tremendous CO. CO can cause harmful health effects by reducing oxygen delivery to the body's organs (like the heart and brain) and tissues. At extremely high levels, CO can cause death.

2.1.4 Nitrogen Oxides

Nitrogen dioxide (NO_2) is one of a group of highly reactive gases known as "oxides of nitrogen," or "nitrogen oxides (NO_x)." Other nitrogen oxides include nitrous acid and nitric acid. EPA's National Ambient Air Quality Standard uses NO_2 as the indicator for the larger group of nitrogen oxides. NO_2 forms quickly from emissions from cars, trucks and buses, power plants, and off-road equipment. In addition to contributing to the formation of ground-level ozone, and fine particle pollution, NO_2 is linked with a number of adverse effects on the respiratory system.

EPA first set standards for NO_2 in 1971, setting both a primary standard (to protect health) and a secondary standard (to protect the public welfare) at 53 parts per billion (ppb), averaged annually. The Agency has reviewed the standards twice since that time, but chose not to revise the annual standards at the conclusion of each review. In January 2010, EPA established an additional primary standard at 100 ppb, averaged over one hour. Together the primary standards protect public

health, including the health of sensitive populations—people with asthma, children, and the elderly. No area of the country has been found to be out of compliance with the current NO_2 standards.

2.1.5 Sulfur Dioxide

Sulfur dioxide (SO_2) is one of a group of highly reactive gases known as "oxides of sulfur." The largest sources of SO_2 emissions are from fossil fuel combustion at power plants (73 %) and other industrial facilities (20 %) as well as Volcanic emissions. Smaller sources of SO_2 emissions include industrial processes such as extracting metal from ore, and the burning of high sulfur-containing fuels by locomotives, large ships, and non-road equipments. SO_2 is linked with a number of adverse effects on the respiratory system.

EPA first set standards for SO_2 in 1971. EPA set a 24-h primary standard at 140 ppb and an annual average standard at 30 ppb (to protect health). EPA also set a 3-h average secondary standard at 500 ppb (to protect the public welfare). In 1996, EPA reviewed the SO_2 NAAQS and chose not to revise the standards.

2.1.6 Lead

Lead is a metal found naturally in the environment as well as in manufactured products. The major sources of lead emissions have historically been from fuels in on-road motor vehicles (such as cars and trucks) and industrial sources. Old painting before 1978 can contain harmful lead. As a result of EPA's regulatory efforts to remove lead from on-road motor vehicle gasoline, emissions of lead from the transportation sector dramatically declined by 95 % between 1980 and 1999, and levels of lead in the air decreased by 94 % between 1980 and 1999. Today, the highest levels of lead in air are usually found near lead smelters. The major sources of lead emissions to the air today are ore and metals processing and piston engine aircraft operating on leaded aviation gasoline.

3 Satellite Data for Studying Air Quality

Satellite remote sensing of air quality has been evolved over last decades. Although the space-borne sensor cannot directly measure the particulate matters or aerosols and trace gases such as CO, NO_2, O_3, and SO_2, we are able to analyze scattered and emitted radiation to accurately derive aerosols and trace gases. The space-borne instruments can be divided into active and passive sensors. The active-sensed instruments, for example, Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) launched on April 28, 2006, sent out signals and received the backscattered signals. Using the backscattered signals, one may be able to

determine aerosol type and mass concentrations (Liu et al. 2009). Most of space-borne remote sensing sensors are passive, which receive from atmosphere/surface scattered and emitted radiation. The visible channels of POLDER, MODIS (Liang et al. 2006), and VIIRS (Cao et al. 2013) have been successfully used to derive aerosol optical depths over dark surfaces. SeaWiFS was for ocean color products. The data are found very valuable to obtain aerosol optical depth over oceans. The ultraviolet sensors like Global Ozone Monitoring Experiment-2 (GOME-2), ozone monitoring instrument (OMI), and ozone mapping profiler suite (OMPS) (Wu et al. 2014) are sensitive to aerosols in particular to absorbing aerosols, although the primary application of those UV sensors is for ozone retrievals. Many passive sensors have been used to derive the products of chemical compositions in the atmosphere. The Earth Observing System (EOS) Microwave Limb Sounder (MLS) on the NASA's EOS Aura satellite, launched on July 15, 2004, has demonstrated great success in retrieving ozone and other atmospheric compositions. MLS observes thermal microwave emission from Earth's "limb" (the edge of the atmosphere) viewing forward along the Aura spacecraft flight direction, scanning its view from the ground to ~90 km every ~25 s. The instrument for the Measurements of Pollution in the Troposphere (MOPITT) is a payload scientific instrument launched into Earth orbit by NASA on board the Terra satellite on December 18, 1999. MOPITT's near-infrared radiometer at 2.3 and 4.7 μm specific focus is on the distribution, transport, sources, and sinks of carbon monoxide in the troposphere. Carbon monoxide, which is expelled from factories, cars, and forest fires, hinders the atmosphere's natural ability to rid itself of harmful pollutants.

The Global Ozone Monitoring Experiment-2 (GOME-2), a European instrument flying on the MetOp-A series of satellites (Launched on October 19, 2006) was designed by the European Space Agency to measure atmospheric ozone trace gases and ultraviolet radiation http://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/MetopDesign/GOME2/index.html). It is a scanning instrument (scan width 1920 km) with near global coverage daily. The field of view on the ground is 80 km by 40 km. It also provides accurate information on the total column amount of nitrogen dioxide, sulfur dioxide, water vapor, oxygen/oxygen dimmer, bromine oxide, and other trace gases, as well as aerosols.

The Multi-angle Imaging SpectroRadiometer (MISR) is a scientific instrument on the Terra satellite launched by NASA on December 18, 1999. This device is designed to measure the intensity of solar radiation reflected by the Earth system (planetary surface and atmosphere) in various directions and spectral bands; it became operational in February 2000. Data generated by this sensor have been proven useful in a variety of applications including atmospheric sciences, climatology, and monitoring terrestrial processes.

The MISR instrument consists of an innovative configuration of nine separate digital cameras that gather data in four different spectral bands of the solar spectrum. One camera points toward the nadir, while the others provide forward and afterward view angles at 26.1°, 45.6°, 60.0°, and 70.5°. As the instrument flies

overhead, each region of the Earth's surface is successively imaged by all nine cameras in each of four wavelengths (blue, green, red, and near-infrared).

The data gathered by MISR are useful in climatological studies concerning the disposition of the solar radiation flux in the Earth's system. MISR is specifically designed to monitor the monthly, seasonal, and long-term trends of atmospheric aerosol particle concentrations including those formed by natural sources and by human activities, upper air winds and cloud cover, type, height, as well as the characterization of land surface properties, including the structure of vegetation canopies, the distribution of land cover types, or the properties of snow and ice fields, amongst many other biogeophysical variables.

NESDIS GOME 2 (MetOp-A) total ozone products, based on the SBUV/2 version 8 algorithms are produced in binary and BUFR formats. The algorithm also produces aerosol index and reflectivity values, which are included in the total ozone binary product.

Infrared hyperspectral sensors such as AIRS, IASI, and CrIS are useful to retrieve trace gases. IASI samples radiance in a spectral resolution of 0.25 cm^{-1}, providing a high sensitivity to trace gases. CrIS data of a full spectral resolution of 0.625 cm^{-1} can be used to derive CO (Liu and Xiao 2014).

4 Community Radiative Model

CRTM is a sensor-based radiative transfer model. It supports more than 100 sensors including sensors on most meteorological satellites and some from other remote sensing satellites. The CRTM is composed of four important modules for gaseous transmittance, surface emission and reflection, cloud and aerosol absorption and scatterings, and a solver for a radiative transfer. The CRTM was designed to meet users' needs. Many options are available for users to choose: input surface emissivity; select a subset of channels for a given sensor; turn off scattering calculations; compute radiance at aircraft altitudes; compute aerosol optical depth only; and threading of the CRTM. Figure 1 shows the interface diagram for users (public interface) and internal modules for developers contained in the lower dashed box. The CRTM forward model is used to simulate from satellite-measured radiance, which can be used to verify measurement accuracy, uncertainty, and long-term stability. The k-matrix module is used to compute jacobian values (i.e., radiance derivative to geophysical parameters), which is used for the inversion processing in retrieval and radiance assimilations. Using tangent linear and adjoint modules is equivalent to using k-matrix module and is also applied to some application in radiance assimilation. In the following subsessions, we will describe the CRTM modules in detail.

The CRTM is a library for users to link, instead of a graphic user interface. By the CRTM initialization, user selects the sensor/sensors and surface emissivity/reflectance lookup tables. Developers may incorporate their own expertise into the CRTM for any desired applications. The gaseous transmittance describes atmospheric gaseous

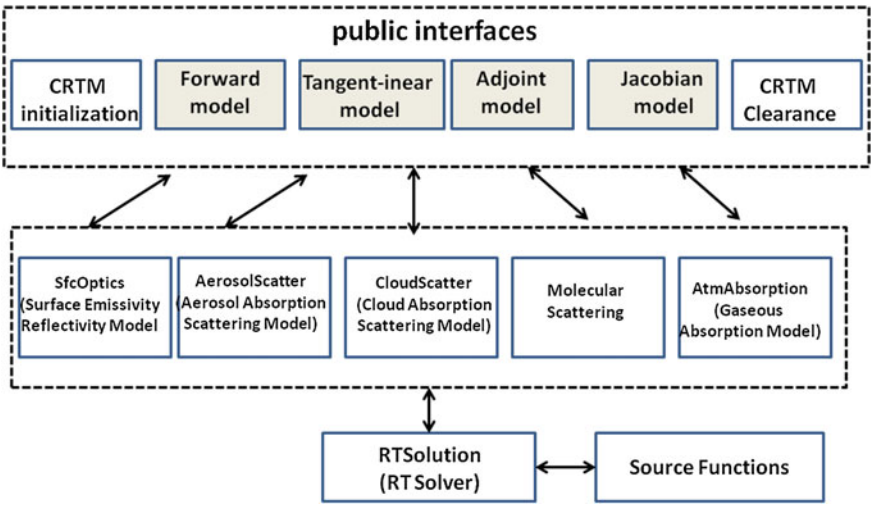


Fig. 1 An interface diagram of the community radiative transfer model. The modules in the public interfaces (*upper dashed box*) are accessed by users. The modules below the *upper dashed box* are for developers

absorption, so that one can utilize remote sensing information in data assimilation/retrieval systems for atmospheric temperature, moisture, and trace gases such as CO_2, O_3, N_2O, CO, and CH_4 (Chen et al. 2012). The aerosol module is fundamental to acquire aerosol type and concentration for studying air quality. The cloud module contains optical properties of six cloud types, providing radiative forcing information for weather forecasting and climate studies. The CRTM surface model includes surface static and atlas-based emissivity/reflectivity for various surface types. Two radiative solutions have been implemented into the CRTM. The advanced doubling-adding (ADA) method (Liu and Weng 2013, 2006) is chosen as a baseline. The successive order of interaction (SOI) radiative transfer model (Heidinger et al. 2006) developed at the University of Wisconsin, has also been implemented in the CRTM.

For a new sensor, the CRTM team can generate spectral and transmittance coefficient files as long as the spectral response data of the new sensor is available. Once the spectral and transmittance coefficient files are created, the CRTM is ready for the new sensor. The new surface emissivity model may be supplied if the user wants to derive surface emissivity for the new sensor. The CRTM user interface provides forward, tangent linear, adjoint, and k-matrix functions to compute radiance (also microwave and infrared brightness temperature) and sensitivities of radiance to atmospheric/surface parameters. The NOAA Microwave Integrated Retrieval System (MiRS) (Boukabara et al. 2007) and NCEP data assimilation system use the k-matrix. The Weather and Research Forecasting (WRF) model uses the tangent linear and adjoint models. The NOAA Integrated Calibration/Validation

System Long-term monitoring system (http://www.star.nesdis.noaa.gov/icvs/) uses the forward model to compare the CRTM simulation with satellite measurements.

The validation of radiative transfer calculations is very challenging because it depends on the model assumptions (e.g., spherical or nonspherical scatterers), measurement errors, uncertainities in inputs, and others. Many evaluations of the CRTM model simulations versus observations have been performed. Saunder et al. (Saunders et al. 2007) summarized the brightness temperature differences between the line-by-line (LBL) model and fast radiative transfer models including the CRTM where the Optical Path Transmittance (OPTRAN) algorithm is used. The difference in the standard deviation under clear-sky conditions is generally less than 0.1 K for the Atmospheric Infrared Sounder (AIRS). Under cloudy conditions, the difference for the AIRS is about 0.2 K (Ding et al. 2011). Due to the approximation in cloud scattering calculations, the difference between the LBL model and the CRTM for the High-Resolution Infrared Radiation Sounder/3 (HIRS/3) under cloudy conditions can reach 0.4 K (Liu et al. 2013). It is more difficult to estimate the errors between the CRTM model calculations and measurements. We can only give the estimate from limited applications. Liang et al. (2009) implemented the CRTM into the NOAA Monitoring System of IR Clear-sky Radiances over Oceans. The difference between the CRTM simulations and NOAA *Advanced Very High Resolution Radiometer* (AVHRR) observations is about 0.5 K. For the Stratosphere Sounding Unit (SSU), the difference is about 1 K (Liu and Weng 2009). Han et al. (2007) compared the CRTM simulations with the *Special Sensor Microwave Imager/Sounder* (*SSMIS*) measurements and found both agreed within about 2 K. NOAA NCEP (http://www.emc.ncep.noaa.gov/gmb/gdas/radiance/esafford/wopr/index.html) monitors the difference under clear-sky conditions for more than 20 sensors and found that the difference is generally less than 1 K. Under cloudy conditions, uncertainties in inputs are the main source, resulting in a large difference. The difference is about 2 K for microwave sounding channels and 4 K for microwave window channels (Chen et al. 2008).

4.1 Radiative Transfer Equation and Solver

The CRTM is one-dimensional radiative transfer model. This implies that atmosphere is assumed homogeneous in the horizontal direction, so-called a plane-parallel atmosphere. For the plane-parallel atmosphere, a vector radiative transfer model can be written as

$$\mu \frac{\mathrm{d}\mathbf{I}(\tau,\Omega)}{\mathrm{d}\tau} = -\mathbf{I}(\tau,\Omega) + \frac{\varpi}{4\pi}\int_0^{2\pi}\int_{-1}^{1} \mathbf{M}(\tau;\Omega;\Omega')\mathbf{I}(\tau,\Omega')\mathrm{d}\Omega' + \mathbf{S}(\tau,\Omega;\Omega_0), \quad (1a)$$

where $\mathbf{\Omega}$ represents a beam for a pair (μ, ϕ) in an incoming or an outgoing directionor

$$\mu \frac{\mathrm{d}\mathbf{I}(\tau,\mu,\phi)}{\mathrm{d}\tau} = -\mathbf{I}(\tau,\mu,\phi) + \frac{\varpi}{4\pi}\int_0^{2\pi}\int_{-1}^{1} \mathbf{M}(\tau;\mu,\phi;\mu',\phi')\mathbf{I}(\tau,\mu',\phi')d\mu' d\phi' + \mathbf{S}(\tau,\mu,\phi;\mu_0,\phi_0) \tag{1b}$$

and

$$S = (1-\varpi)B[T(\tau)]\begin{bmatrix}1\\0\\0\\0\end{bmatrix} + \frac{\varpi F_0}{4\pi}\exp(-\tau/\mu_0)\begin{bmatrix}M_{11}(\mu,\phi;\mu_0,\phi_0)\\M_{12}(\mu,\phi;\mu_0,\phi_0)\\M_{13}(\mu,\phi;\mu_0,\phi_0)\\M_{14}(\mu,\varphi;\mu_0,\varphi_0)\end{bmatrix}, \tag{1c}$$

where $\mathbf{M}$ is the phase matrix; $\mathbf{I} = [I, Q, U, V]^{\mathrm{T}}$; $B(T)$ the Planck function at a temperature T; F_0 the solar spectral constant; μ_0 the cosine of sun zenith angle; ϖ the single-scattering albedo; and τ the optical thickness.

Equation (1a) can be solved by a standard discrete method for beams rather than separate zenith angles from azimuthal angles. The advantage of this approach is that one can directly use the phase function without doing any truncation. The disadvantage is that it demands more memory for a large number of beams that can slow the computation. The approach can be useful for more accurate simulations in the future when computational capacity increases.

Equation (1b) is commonly used in radiative transfer models including the CRTM. The equation can be solved by some standard routines such as the multilayer discrete ordinate method (Stamnes et al. 1988), the Doubling-adding method (Evans and Stephens 1991), and the matrix operator method (Liu and Ruprecht 1996). Essentially, the azimuthal dependence of Stokes vector is expanded into a series of Fourier harmonics. The amplitude of each Fourier component is a function of zenith angles. Furthermore, the amplitude is discretized at a series of zenith angles (or streams) so that the combined Stokes cosine and sine harmonics can be simplified as

$$\begin{aligned}\mu_i \frac{\mathrm{d}}{\mathrm{d}\tau}\begin{bmatrix}\mathbf{I}_m(\tau,\mu_i)\\-\mathbf{I}_m(\tau,\mu_{-i})\end{bmatrix} &= \begin{bmatrix}\mathbf{I}_m(\tau,\mu_i)\\\mathbf{I}_m(\tau,\mu_{-i})\end{bmatrix}\\ &\quad - \varpi\sum_{j=1}^{N}\begin{bmatrix}\mathbf{M}_m(\mu_i,\mu_j) & \mathbf{M}_m(\mu_i,\mu_{-j})\\\mathbf{M}_m(\mu_{-i},\mu_j) & \mathbf{M}_m(\mu_{-i},\mu_{-j})\end{bmatrix}\begin{bmatrix}\mathbf{I}_m(\tau,\mu_j)\\\mathbf{I}_m(\tau,\mu_{-j})\end{bmatrix}w_j\\ &\quad - \begin{bmatrix}\mathbf{S}_m(\tau,\mu_i,\mu_0)\\\mathbf{S}_m(\tau,\mu_{-i},\mu_0)\end{bmatrix}.\end{aligned} \tag{2}$$

where μ_i and w_i are Gaussian quadrature points and weights, respectively. Note that $\mu_{-i} = -\mu_i$ and $w_{-i} = w_i$. According to the properties of the phase matrix, the radiance components at sinusoidal and cosinusoidal modes can be decoupled, recombined, and solved independently (Weng and Liu 2003).

The summation in Eq. (2) is over discrete zenith angles or steams. For $N = 1$ in Eq. (2), it is called two-stream radiative transfer to represent one upward and one downward directions. It is called eight-stream radiative transfer when $N = 4$. The computational time is proportional to N^3.

In most radiative transfer schemes, the phase function can be expanded in terms of Legendre polynomials as follows:

$$P(\cos\Theta) = \sum_{l=0}^{M} \omega_l P_l(\cos\Theta) \tag{3}$$

where Θ is the scattering angle, $P_l(\cos\Theta)$ are Legendre polynomials, ω_l are the expansion coefficients, and the value of the upper summation limit M should be equal or smaller than N used in Eq. (2).

Thousands of terms may be required for accurate phase function of an ice cloud in the visible spectrum, but the computation burden makes this an unattractive approach. One way of reducing the necessary Legendre terms is by truncating the forward scattering peak and renormalizing the phase function. Wiscombe (1957) proposed a δ–M method to truncate Legendre polynomials. The δ–M method can effectively remove the forward peak in a rigorous mathematical way, but it may cause large spikes in the phase function for small scattering angles (<20°). The δ-fit method (Hu et al. 2000) truncates the forward scattering peak in the phase function and more optimally fits the remaining phase function by selecting a set of optimized expansion coefficients in Eq. (3). We applied the δ-fit method in this study.

To achieve the same accuracy in multiple scattering calculations with the truncated phase functions as with the nontruncated phase functions, an adjustment must be made to the optical thickness and single-scattering albedo via the following relations (Liou 2002):

$$\tau' = (1 - f\omega)\tau, \tag{4}$$

$$\omega' = \frac{(1-f)\omega}{1 - f\omega}, \tag{5}$$

where f indicates the portion of the scattered energy associated with the truncated forward peak. In Eqs. (4) and (5), τ and ω are the original optical thickness and single-scattering albedo, respectively, whereas τ' and ω' are the optical thickness and single-scattering albedo associated with the truncated phase function.

After we have the phase matrix, Eq. (2) for each harmonic component can be expressed as

$$\frac{\mathrm{d}\mathbf{I}}{\mathrm{d}\tau} = \mathbf{A}\mathbf{I} - \mathbf{S}, \tag{6}$$

where

$$\mathbf{I} = [\mathbf{I}(\tau, \mu_1), \mathbf{I}(\tau, \mu_2), \ldots, \mathbf{I}(\tau, \mu_N), \mathbf{I}(\tau, \mu_{-1}), \mathbf{I}(\tau, \mu_{-2}), \ldots, \mathbf{I}(\tau, \mu_{-N})]^{\mathrm{T}}, \tag{7}$$

$$S = (1 - \varpi)B(T)\delta_{m0}\begin{bmatrix} \mathbf{u}^{-1} & \mathbf{0} \\ \mathbf{0} & -\mathbf{u}^{-1} \end{bmatrix}\boldsymbol{\Xi} + \frac{\varpi F_0}{\pi}\exp(-\tau/\mu_0)\boldsymbol{\Psi}, \tag{8}$$

where $\mathbf{u}$ is a $4N$ by $4N$ matrix that has nonzero elements at its diagonal direction such as

$$u = [\mu_1, \mu_1, \mu_1, \mu_1, \mu_2, \ldots, \mu_N, \mu_N, \mu_N, \mu_N]_{\text{diagnonal}}, \tag{9a}$$

$\boldsymbol{\Xi}$ and $\boldsymbol{\Psi}$ are vectors that have $8N$ elements as

$$\boldsymbol{\Xi} = [1, 0, 0, 0, 1, 0, 0, 0, \ldots, 1, 0, 0, 0]^{\mathrm{T}}; \tag{9b}$$

and

$$\begin{aligned}\boldsymbol{\Psi} = [&M_{11}(\mu_1, \mu_0)/\mu_1, M_{12}(\mu_1, \mu_0)/\mu_1, M_{13}(\mu_1, \mu_0)/\mu_1, M_{14}(\mu_1, \mu_0)/\mu_1, \\ &M_{11}(\mu_2, \mu_0)/\mu_2, M_{12}(\mu_2, \mu_0)/\mu_2, M_{13}(\mu_2, \mu_0)/\mu_2, M_{14}(\mu_2, \mu_0)/\mu_2, \ldots, \\ &M_{11}(\mu_{-N}, \mu_0)/\mu_{-N}, M_{12}(\mu_{-N}, \mu_0)/\mu_{-N}, M_{13}(\mu_{-N}, \mu_0)/\mu_{-N}, M_{14}(\mu_{-N}, \mu_0)/\mu_{-N}],\end{aligned} \tag{9c}$$

and the composite phase matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{u}^{-1} & 0 \\ 0 & -\mathbf{u}^{-1} \end{bmatrix}\begin{bmatrix} \mathbf{E} - \varpi\,\mathbf{M}(u, u) & \varpi\,\mathbf{M}(u, -u) \\ \varpi\,\mathbf{M}(-u, u) & \mathbf{E} - \varpi\,\mathbf{M}(-u, -u) \end{bmatrix} = \begin{bmatrix} \boldsymbol{\alpha}_1 & \boldsymbol{\beta}_1 \\ -\boldsymbol{\beta}_2 & -\boldsymbol{\alpha}_2 \end{bmatrix} \tag{9d}$$

where $\mathbf{E}$ is a unit matrix. For a pair of zenith angles (μ_i, μ_j), both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are $4N$ by $4N$ matrices and are related to the elements of the phase matrices as

$$\boldsymbol{\alpha}_1(\mu_i, \mu_j) = [\mathbf{E} - \varpi\,\mathbf{M}_m(\mu_i, \mu_j)]/\mu_i \tag{10a}$$

$$\boldsymbol{\beta}_1(\mu_i, \mu_{-j}) = \varpi\,\mathbf{M}_m(\mu_i, \mu_{-j})/\mu_i \tag{10b}$$

$$\boldsymbol{\alpha}_2(\mu_{-i}, \mu_{-j}) = [\mathbf{E} - \varpi\,\mathbf{M}_m(\mu_{-i}, \mu_{-j})]/\mu_i \tag{10c}$$

$$\boldsymbol{\beta}_2(\mu_{-i}, \mu_j) = \varpi\,\mathbf{M}_m(\mu_{-i}, \mu_j)/\mu_i \tag{10d}$$

Equations (1b) to (10a–10d) are a common approach needed by radiative solvers. The equations can be solved using double-adding (Evans and Stephens 1991), matrix operator method (MOM) (Liu and Weng 2013; Liu and Ruprecht 1996), VDISORT (Weng and Liu 2003), Successive Order of Interaction (SOI) Radiative Transfer Model (Heidinger et al. 2006), and advanced double-adding (ADA) method. In the following, we discuss the ADA for radiative intensity only, a scalar radiative transfer model. The ADA is a default solver in the CRTM. SOI can be selected by users. Both thermal emission source and solar reflection parts are used operationally in supporting of radiance assimilation and satellite products generations. The thermal emission part was documented in the paper (Liu and Weng 2006). The description of the solar reflection part will be added here. Equation (6) for intensity only can be rewritten in a matrix–vector form as

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\begin{bmatrix}\mathbf{I_u}\\ \mathbf{I_d}\end{bmatrix} = -\begin{bmatrix}\boldsymbol{\alpha} & \boldsymbol{\beta}\\ -\boldsymbol{\beta} & -\boldsymbol{\alpha}\end{bmatrix}\begin{bmatrix}\mathbf{I_u}\\ \mathbf{I_d}\end{bmatrix} - (1-\varpi)B(T)\begin{bmatrix}\mathbf{u}^{-1}\boldsymbol{\Xi}\\ -\mathbf{u}^{-1}\boldsymbol{\Xi}\end{bmatrix} + \frac{\varpi F_0}{\pi}\exp(-\tau/\mu_0)\begin{bmatrix}\mathbf{u}^{-1}\boldsymbol{\Phi_u}\\ -\mathbf{u}^{-1}\boldsymbol{\Phi_d}\end{bmatrix} \tag{11}$$

where $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are N by N matrices (All bold letters and symbols indicate either matrix or vector.) and

$$\boldsymbol{\alpha}(\mu_i,\mu_j) = [\varpi P(\mu_i,\mu_j)w_j - \delta_{ij}]/\mu_i, \tag{12a}$$

$$\boldsymbol{\beta}(\mu_i,\mu_{-j}) = \varpi P(\mu_i,\mu_{-j})w_j/\mu_i, \tag{12b}$$

δ_{ij} is the Kronecker delta. The subscripts u and d indicate upward and downward directions, respectively. $\mathbf{u}$ is an N by N matrix that has nonzero elements in its diagonal such as

$$\mathbf{u} = [\mu_1,\mu_2,\ldots,\mu_N]_{\text{diagnonal}}, \tag{12c}$$

$\boldsymbol{\Xi}$ is a vector of N elements as

$$\boldsymbol{\Xi} = [1,1,\ldots,1]^{\mathrm{T}}. \tag{12d}$$

For an infinitesimal optical depth δ_0, multiple scattering can be neglected and the reflection matrix can be expressed as (Plass et al. 1973)

$$\mathbf{r}(\delta_0) = \delta_0\boldsymbol{\beta}, \tag{12e}$$

and the transmission matrix can be written as

$$\mathbf{t}(\delta_0) = \mathbf{E} + \boldsymbol{\alpha}\delta_0, \tag{12f}$$

$\mathbf{E}$ is an N by N unit matrix.

Using the doubling procedure from Van de Hulst (1963), the reflection and transmission matrices for a finite optical depth ($\delta = \delta_n = 2^n\delta_0$) can be computed by doubling the optical depth (i.e., $\delta_{i+1}/\delta_i = 2$) recursively:

$$\mathbf{r}(\delta_{i+1}) = \mathbf{t}(\delta_i)[\mathbf{E} - \mathbf{r}(\delta_i)\mathbf{r}(\delta_i)]^{-1}\mathbf{r}(\delta_i)\mathbf{t}(\delta_i) + \mathbf{r}(\delta_i) \tag{13a}$$

and

$$\mathbf{t}(\delta_{i+1}) = \mathbf{t}(\delta_i)[\mathbf{E} - \mathbf{r}(\delta_i)\mathbf{r}(\delta_i)]^{-1}\mathbf{t}(\delta_i) \tag{13b}$$

for $i = 0, n - 1$. We denote $\mathbf{r}(k) = \mathbf{r}(\delta_n)$ and $\mathbf{t}(k) = \mathbf{t}(\delta_n)$ for the reflection and transmission matrices of kth layer.

There exist formulas for building the layer source functions (Heidinger et al. 2006) depending on the Planck function of the temperature at the top of the layer and the gradient of Planck function over the layer optical depth. However, the formulas are complicated and computationally expensive. In this study, we found a very simple and strict expression for the layer source function using the existing layer reflection and transmission matrices. For an atmospheric layer of an optical depth δ and having the top temperature of T_1 and the bottom temperature of T_2, the upward layer source function can be derived as (see Appendix A of (Liu and Weng 2006)

$$\begin{aligned}\mathbf{S_u} = {} & [(\mathbf{E} - \mathbf{t} - \mathbf{r})B(T_1) - (B(T_2) - B(T_1))\mathbf{t} + \frac{B(T_2) - B(T_1)}{(1 - \varpi g)\delta}(\mathbf{E} + \mathbf{r} - \mathbf{t})\mathbf{u}]\boldsymbol{\Xi} \\ & + \frac{\omega F_0}{\pi}\exp\left(-\frac{\tau}{\mu_0}\right)[(\mathbf{E} - \mathbf{t}\exp\left(\exp\left(-\frac{\delta}{\mu_0}\right)\right)\boldsymbol{\Psi}_\mathbf{u} - \mathbf{r}\boldsymbol{\Psi}_\mathbf{d}]\end{aligned} \tag{14a}$$

and the downward source of the layer can be written as

$$\begin{aligned}\mathbf{S_d} = {} & [(\mathbf{E} - \mathbf{t} - \mathbf{r})B(T_1) + (B(T_2) - B(T_1))(\mathbf{E} - \mathbf{r}) + \frac{B(T_2) - B(T_1)}{(1 - \varpi g)\delta}(\mathbf{t} - \mathbf{E} - \mathbf{r})\mathbf{u}]\boldsymbol{\Xi} \\ & + \frac{\omega F_0}{\pi}\exp\left(-\frac{\tau}{\mu_0}\right)\left[\left(\exp\left(-\frac{\delta}{\mu_0}\right)\mathbf{E} - \mathbf{t}\right)\boldsymbol{\Psi}_\mathbf{d} - \mathbf{r}\exp\left(-\frac{\delta}{\mu_0}\right)\boldsymbol{\Psi}_\mathbf{u}\right],\end{aligned} \tag{14b}$$

where ϖ and g are single-scattering albedo and asymmetry factor of the layer, respectively. τ is the optical depth from the top of the atmosphere to the top of this current layer.

$$\begin{bmatrix}\boldsymbol{\Psi}_\mathbf{u} \\ \boldsymbol{\Psi}_\mathbf{d}\end{bmatrix} = -\frac{\omega F_0}{(1 + \delta_{0m})}\begin{bmatrix}\boldsymbol{\alpha} + \mathbf{E}/\mu_0 & \boldsymbol{\beta} \\ -\boldsymbol{\beta} & \boldsymbol{\alpha} - \mathbf{E}/\mu_0\end{bmatrix}\begin{bmatrix}\boldsymbol{\Phi}_\mathbf{u} \\ \boldsymbol{\Phi}_\mathbf{d}\end{bmatrix} \tag{15}$$

The new expressions for the layer source functions take a very little extra computation time. The expression can also be applied for other radiative transfer models such as matrix operator method (Fischer and Grassl 1984).

Equations (13a, 13b)–(14a, 14b) give the layer reflection and transmission matrices as well as the source vectors at the upward and downward directions. For planetary atmosphere, the atmosphere may be divided into n optically homogeneous layers. The optical properties (e.g., extinction coefficient, single-scattering albedo, and phase matrix) are the same within each layer although the temperature may vary within the layer. The adding method is for integrating the surface and multiple atmospheric layers. The method was applied to flux calculation using a two-stream approximation. The method was also used in radiance calculations with multiple scatterings using a two-stream approximation (Schmetz and Raschke 1981). In the following, we briefly describe the methodology. We denote $\mathbf{R_u}(k)$ for reflection matrix and $\mathbf{I_u}(k)$ for radiance vector at the level k in the upward direction and $k = n$ and $k = 0$ represent the surface level and the top of the atmosphere, respectively. The adding method starts from surface without atmosphere. At the surface, $\mathbf{R_u}(n)$ is the surface reflection matrix and $\mathbf{I_u}(n)$ equals the surface emissivity vector multiplied by the Planck function at the surface temperature. The upward reflection matrix and radiance at the new level can be obtained by adding one layer from the present level:

$$\mathbf{R}(k-1) = \mathbf{r}(k) + \mathbf{t}(k)\,[\mathbf{E} - \mathbf{R}(k)\,\mathbf{r}(k)]^{-1}\,\mathbf{R}(k)\,\mathbf{t}(k), \tag{16a}$$

$$\begin{aligned}\mathbf{I_u}(k-1) &= \mathbf{S_u}(k) + \mathbf{t}(k)\,[\mathbf{E} - \mathbf{R}(k)\,\mathbf{r}(k)]^{-1}\mathbf{R}(k)\mathbf{S_d}(k) + \mathbf{t}(k)\,[\mathbf{E} - \mathbf{R}(k)\mathbf{r}(k)]^{-1}\mathbf{I_u}(k) \\ &= \mathbf{S_u}(k) + \mathbf{t}(k)[\mathbf{E} - \mathbf{R}(k)\,\mathbf{r}(k)]^{-1}\,[\mathbf{R}(k)\mathbf{S_d}(k) + \mathbf{I_u}(k)].\end{aligned} \tag{16b}$$

The physical meaning of Eq. (16a, 16b) is obvious. The first term on the right side of Eq. (16a) is the reflectance of the layer to be added. The second term on the right side of Eq. (16a) is the reflectance due to the radiation from the new level transmitted to and multiple reflected by the present level and then transmitted back to the new level. The three terms on the right side of Eq. (16b) represent the upward layer source, from the present level reflected layer downward source, and from the present level transmitted upward radiance, respectively. The upward radiance $\mathbf{I_u}$ at the top of the atmosphere can be obtained by looping the index from $k = n$ to $k = 1$ and adding the contribution from cosmic background radiance (Planck function at the temperature of 2.7 K) vector $\mathbf{I}_{\text{sky}}$, that is,

$$\mathbf{I_u} = \mathbf{I_u}(0) + \mathbf{R}_u(0)\ \mathbf{I}_{\text{sky}} \tag{17}$$

Equations (12a–12f)–(17) give the necessary and sufficient formulas for advanced doubling-adding method. It needs to be mentioned that the above procedure is for the upward radiance at the top of the atmosphere. It is sufficient for the satellite data assimilation. However, an additional loop from the top to the surface is necessary in order to obtain the vertical profiles of radiances at both upward and downward directions.

For the viewing angle departure from the angles at Gaussian quadrature points, an additional stream as extra Gaussian quadrature point associated with an integration weight of zero may be inserted to have N quadrature points in total in either

upward or downward direction. For this case, the upward intensity vector will contain the upward solutions at $N - 1$ quadrature points and at a specified viewing angle. The result by inserting the additional stream is exactly the same as inserting the multiple scattering solutions at $N - 1$ quadrature points back to the integration equation for the specified viewing angle [see Eqs. (24) and (25) of Stamnes et al. (Chen et al. 2008)]. However, the present procedure avoids using extra codes for the specified viewing angle so that it much simplifies all forward, tangent linear, and adjoint codings.

All rigorous discrete radiative transfer solvers should be very accurate. The model intercomparisons in the paper (Liu and Weng 2006) are adopted here between the doubling-adding model (Evans and Stephens 1991), VDISORT (Weng and Liu 2003), and the advanced doubling-adding method. We use the CRTM platform which allows us to insert various solvers for radiative transfer calculations. Three solvers mentioned above share the same atmospheric optic data, the same surface emissivity and reflectivity, and the same Planck function for atmosphere, surface, and the cosmic background. The differences of results from the three solvers are purely from the differences in the solvers. For 24,000 simulations with various clear and cloudy cases, computation times on our personal computer are 1041, 29, and 17 s for DA, VDISORT, and ADA models, respectively. ADA is about 1.7 times faster than VDISORT and 61 times faster than DA. The huge gain of ADA to DA is partly contributed by the efficiency of matrix and vector manipulation in FORTRAN 95 because DA code is still in FORTRAN 77. The maximum difference of the simulated brightness temperatures between using the three solvers for AMSU-A channels and 281 selected Atmospheric InfraRed Sounder (AIRS) channels is less than 0.01 K. The subset of AIRS data used in NCEP data assimilation contains necessary information on atmospheric temperature and water vapor (Goldberg et al. 2003). Tables 2 and 3 list the comparison of the brightness temperatures for AMSU-A water vapor channel at 23.8 GHz and the infrared window channel of AIRS at 10.88 μm computed from ADA, VDISORT, DA methods, respectively. A profile containing temperature and water vapor as well as ozone from our test dataset for OPTRAN is selected. For the microwave calculation, a rain cloud having an effective particle size of 200 μm and 0.5 mm rain water content was put at 850 hPa. One layer ice cloud having the same effective particle size and 0.1 mm ice water path is located at 300 hPa. A wind speed of 5 ms^{-1} over ocean is used. The maximum difference of the brightness temperature computed from the three models is less than 0.01 K (see Table 2). For the infrared calculation, an ice cloud having an effective particle size of 20 μm and 0.1 mm ice water path was located at 300 hPa and a liquid water cloud at 850 hPa having an effective particle size of 10 μm and 0.5 mm are chosen. The results computed from the three models agree very well (see Table 3).

Table 2 Comparison of brightness temperatures at 23.8 GHz (AMSU-A channel 1) computed from advanced doubling-adding method (ADA), VDISORT, and doubling-adding method (DA)

Zenith angle	ADA	VDISORT	DA
0	272.9645	272.9656	272.9655
10	272.9358	272.9369	272.9369
20	272.8342	272.8354	272.8354
30	272.6054	272.6065	272.6064
40	272.0529	272.0542	272.0541
50	271.1577	271.1594	271.1593
65	269.0612	269.0637	269.0635

A rain cloud having an effective radius of 200 μm and 0.5 mm water content was put at 850 hPa. One layer ice cloud having the same effective particle size and 0.1 mm ice water path is located at 300 hPa

Table 3 Comparison of brightness temperatures at 10.88 μm (AIRS channel 256) computed from advanced doubling-adding method (ADA), VDISORT, and doubling-adding method (DA)

Zenith angle	ADA	VDISORT	DA
0	240.7513	240.7514	240.7514
10	240.5512	240.5513	240.5512
20	239.9758	239.9757	239.9757
30	239.1067	239.1065	239.1065
40	238.0799	238.0798	238.0798
50	237.0585	237.0585	237.0585
65	235.6128	235.6129	235.6129

An ice cloud having an effective particle size of 20 μm and 0.1 mm ice water path was located at 300 hPa and a liquid water cloud at 850 hPa having an effective particle size of 10 μm and 0.5 mm are chosen

4.2 Atmospheric Transmittance Models

Line-by-line transmittance model performs for each absorption line. The line width is typically about 10^{-4} cm^{-1} in stratosphere and 10^{-2} cm^{-1} in low troposphere. The line-by-line calculation demands tremendous computations that are not feasible in daily operations. The model often serves as a reference for fast radiative transfer models. LBLrtm (Clough et al. 2005) is one of the best rigorous radiative transfer models for us to generate regression coefficients in the CRTM fast transmittance models. There are two fast transmittance algorithms available in the CRTM: Optical Depth in Absorber Space (ODAS) and Optical Depth in Pressure Space (ODPS); both algorithms are regression-based and differ primarily in vertical coordinates and the application of constraints to smooth vertical structures of the regression coefficients (Chen et al. 2012). The CRTM transmittance coefficients are derived by applying regression algorithms (McMillin et al. 1995) using the line-by-line (LBL) transmittances convolved with the instrument spectral response functions (SRFs) as predictands, and atmospheric state variables as predictors. The CRTM transmittance has two parts: variable gases and fixed gases. For ODAS, H_2O and O_3 are

variable gases and all other gases are treated as fixed gases. For ODPS, six gases (H_2O, CO_2, O_3, N_2O, CO, and CH_4) are variable gases and rest gases are treated as fixed gases. H_2O and O_3 are mandate inputs for both ODAS and ODPS algorithms. CO_2, N_2O, CO, and CH_4 are optional and default values from the ODPS reference profile will be used when inputs are absent. There are two specific transmittance models in the CRTM. The stratospheric sounder unit (SSU) transmittance model (Liu and Weng 2009) was developed to take the CO_2 cell pressure into account. Zeeman effect is considered for microwave sounding channels that are sensitive to upper stratosphere (Han et al. 2007). Figure 2 is the measured (forest green color)

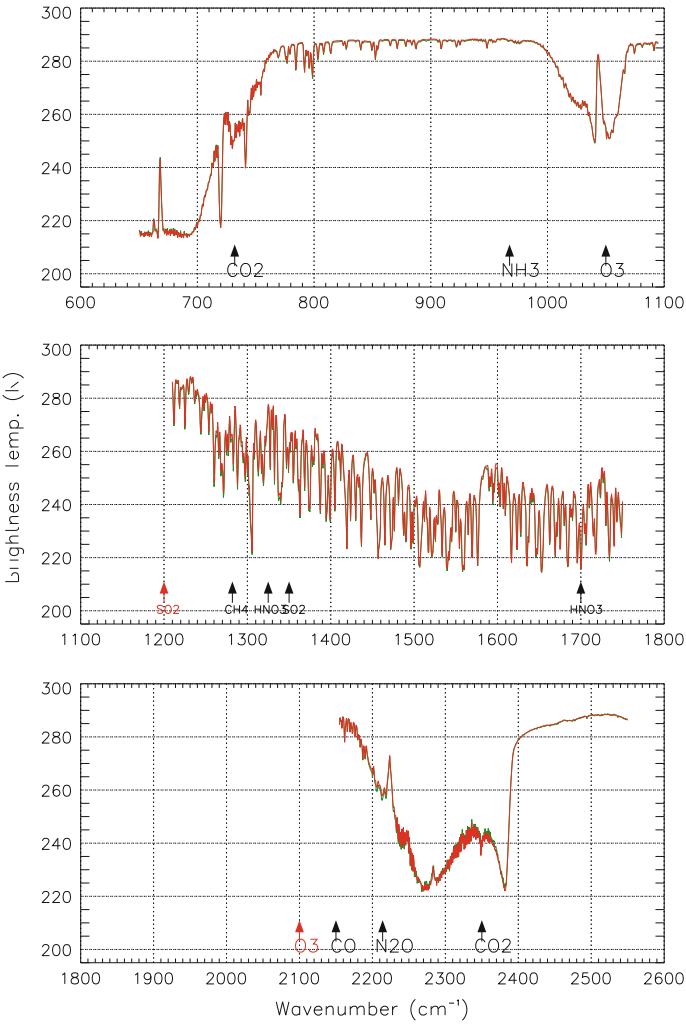


Fig. 2 Measured (*red line*) and the CRTM simulated (forest *green line*) CrIS brightness temperatures. The *upward arrows* indicate the spectral absorption of gases

and the CRTM simulated (red) CrIS brightness temperature of a full spectral resolution over ocean under a clear-sky condition. In the CRTM calculation, radiosonde data and ozonesonde data from the surface to 10 hPa, and from ship measured skin temperatures are used. Above 10 hPa, the ECMWF analysis data are used. Because there was a dry bias in radiosonde water vapor in the upper troposphere during the daytime, we use the ECMWF data for water vapor above 300 hPa. We use the mean value of the ECMWF and ship measured skin temperatures as the surface temperature. Other trace gases such as CO_2, N_2O, CH_4, and CO are taken from a standard tropic atmosphere. The CrIS brightness temperatures are calculated from Hamming apodized radiances at 2211 channels. In general, both measured and simulated CrIS brightness temperatures agree very well. Figure 1 shows main absorption of gases. CrIS observations can be used to retrieve H_2O, O_3, CO_2, N_2O, CO, CH_4, NH_3, SO_2, and HNO_3.

There is a spectral gap between long-wave band and middle-wave band, as well as a gap between mid-wave band and shortwave band. The spectral information there can help the retrieval of SO_2 and volcanic ash.

4.3 Scattering Properties

The scattering of electromagnetic wave by atmospheric particles depends on dielectric constant of particles, particle size distributions, and particle shapes and orientations. In the CRTM, we deal with spherical particles and nonspherical particles with a random orientation. The assumption is common and can simplify the optical properties to a scalar extinction and scattering coefficients and to four phase function elements for spheres and six phase function elements for nonspherical particles with a random orientation. Dielectric constant depends on electromagnetic wavelength and may also depend on temperature. Electromagnetic spectrum of sensors onboard satellites covers ultraviolet (0.20–0.38 μm), visible (violet (0.38–0.45 μm), blue (0.45–0.495 μm), green (0.495–0.570 μm), yellow (0.57–0.59 μm), orange (0.59–0.62 μm), red (0.62–0.75 μm), near-infrared (0.75–3.0 μm), infrared (3.0–50.0 μm), and microwave (1000–30,000 μm) ranges. This partition of the spectrum is approximate since different communities may have own definition (https://en.wikipedia.org/wiki/Infrared).

The ratio of particle diameter to electromagnetic wavelength is one of the key parameters determining scattering. The diameter of a molecule is about 10^{-4} μm. The diameter of aerosols may change from 0.1 to 10 μm. The diameter of clouds varies from 10 to 10,000 μm. Once the dielectric constant and particle size distribution are known, we can use Mie code to compute optical properties such as extinction coefficient, single-scattering albedo, and phase function for spherical particles.

4.4 Molecules

The diameter of molecules in the atmosphere is much smaller than the electromagnetic wavelength in our studies. The scattering by molecules is referred as the Rayleigh scattering. The optical properties of molecular scattering are referred to the literature (Mishchenko et al. 2006).

4.5 Aerosols

Aerosols are small solid or liquid particles in the atmosphere, which can be classified into primary and secondary groups. Primary aerosols, including elemental carbon, organic carbon, sea salt, and mineral dust, are emitted directly from anthropogenic and natural sources, while secondary aerosols, such as sulfate, nitrate, ammonia, and secondary organic carbon, are formed through photochemical, gaseous, aqueous, and heterogeneous chemical reactions in the atmosphere. Aerosols can influence climate directly by scattering and absorbing solar or infrared radiation. They also have impact on climate indirectly by serving as cloud condensation nuclei and ice nuclei and consequently changing albedo and lifetime of clouds.

There exist complex interactions between chemistry, aerosols, and climate. Aerosols are formed through different chemical reactions, and in turn they have impact on photochemical reactions and involve in heterogeneous chemical reactions. Aerosols can modify the global energy balance and lead to climate change, while climate change can affect concentrations of gas phase species and aerosols by changing oxidation capacity of the atmosphere as well as processes of transport, diffusion, deposition, scavenging, and mixing. The chemistry–aerosol–climate interactions play important roles in air pollution, climate change, and water cycle. There are big challenges in understanding the mechanisms of such interactions and in representing aerosols in global and regional climate models. Satellite measurements, in particular multisensor and multichannel measurements, provide useful information about aerosol types and concentrations, which may determine aerosol sources/plumes for regional and weather air quality forecasting. Current satellite aerosol optical depth (AOD) product is helpful, but not accurate to predict aerosol in boundary layer, which is the most important layer to air quality where human breath.

Due to different requirements and heritage from various communities, the calculations of optical properties are not unique, which leads to nonunique results in retrievals and assimilation. We may roughly divide the community as model groups and remote sensing groups. The Goddard Chemistry Aerosol Radiation and Transport (GOCART) model (Chin et al. 2002) is mainly used in a global model. CMAQ focus more on regional area. MODIS-like aerosol is often used in remote sensing group. The CRTM has the optical table for GOCART model for the

spectrum from ultraviolet to infrared. The effect of aerosols on microwave sensors is not considered yet since aerosol size is generally much smaller than microwave length. The optical table for CMAQ is not finalized yet. In this paper, we discuss mainly on the GOCART model.

The GOCART model simulates major tropospheric aerosol components, including sulfate, dust, black carbon (BC), organic carbon (OC), and sea salt aerosols. The sea salt has four sub-types from small particle to large particle. Anthropogenic emission of sulfate (SO_2) from fossil fuel and biofuel combustions and transportations is taken from the Emission Data Base for Global Atmospheric Research (EDGAR) (Olivier et al. 1994). Biomass burning emission of SO_2 is scaled to the seasonal variations of burned biomass data (Duncan et al. 2003). Volcanic emission of SO_2 is from continuously erupting volcanoes and sporatically erupting volcanoes (when data available). Oceanic emission of dimethyl sulfide (DMS) is calculated based on the surface seawater concentrations of DMS and 10-m winds over the ocean using an empirical formula (Liss and Merlivat 1986). Dust particles ranging from 0.1 to 10 μm in radius are considered in the model with eight size groups (0.1–0.18, 0.18–0.3, 0.3–0.6, 0.6–1, 1–1.8, 1.8–3, 3–6, and 6–10 μm). The biomass burning emissions of OC and BC are estimated from the database of seasonal and interannual variations in the burned biomass (Duncan et al. 2003), developed from long-term satellite observations of global fire counts and aerosol index and an annual mean burned biomass inventory. Anthropogenic emissions of OC and BC are taken from a global dataset (Cooke et al. 1999). In addition to direct emissions, the production of OC from terrestrial source is estimated from the emission of volatile organic compounds (Guenther et al. 1995). Sea salt emission from the ocean is highly dependent on the surface wind speed.

Aerosol particle size distribution plays an important role in calculating the optical properties of aerosols. In this study, we assumed a spherical aerosol particle. The aerosol size distribution may have multiple modes. For each mode, a typical aerosol size function is assumed to be the lognormal distribution (d'Almeida 1991; Han et al. 2007) for N particles within the mode,

$$n(\ln r) = \frac{N}{\sqrt{2\pi}\ln(\sigma_g)}\exp\left[-\frac{1}{2}\left(\frac{\ln r - \ln r_g}{\ln(\sigma_g)}\right)^2\right], \tag{18}$$

where r is a radium, r_g the geometric median radius, and σ_g the geometric mean standard deviation. The kth moment of the distribution can be expressed as (Binkowski and Roselle 2003)

$$M_k = \int_{-\infty}^{\infty} r^k n(\ln r) d\ \ln(r)\ = r_g^k \exp\left[\frac{k^2}{2}\ln^2(\sigma_g)\right]. \tag{19}$$

M_0 is the number N of aerosol particles. M_2 and M_3 are proportional to the total particulate surface area and volume, respectively. Thus, the effective radius (r_{eff}) can be defined as

$$r_{\text{eff}} = \frac{M_3}{M_2} = r_g \exp\left[\frac{5}{2}\ln^2(\sigma_g)\right]. \tag{20}$$

Table 4 lists the size parameters and computed mass extinction coefficients used in the CRTM. For sulfate, sea salt, hydrophilic OC and BC, water uptake effect needs to be included. The particle size increases as the relative humidity (RH) of ambient atmosphere increases. The reflective index needs to be calculated by considering water content in aerosols. The effective radius growth factor for hygroscopic aerosols may be theoretically calculated or obtained from a precalculated lookup table (Chin et al. 2002). Once the growth factor a_g is evaluated, the refractive index n for the hygroscopic aerosol can be calculated using a volume mixing method as

$$n_r = n_w + (n_o - n_w) \times a_g^3, \tag{21}$$

where n_o and n_w are the refractive indices for dry aerosols and water, respectively. We adopt refractive index n_o from the software package of Optical Properties of Aerosols and Clouds (OPAC) (Hess et al. 1998). The water refractive index is given by Hale and Querry (1973).

Table 4 Goddard Chemistry Aerosol Radiation and Transport (GOCART) aerosol optical properties at 550 nm for dry air

Aerosol type	Density [g cm^{-3}]	Effective radius [μm]	Standard deviation σ [μm]
Sulfate	1.7	0.242	2.03
OC1 (hydrophobic)	1.8	0.087	2.20
OC2 (hydrophilic)	1.8	0.087	2.20
BC1 (hydrophobic)	1	0.036	2.0
BC2 (hydrophilic)	1	0.036	2.0
SeaSalt1 (size range)	2.2	0.3	2.03
SeaSalt2	2.2	1.0	2.03
SeaSalt3	2.2	3.25	2.03
SeaSalt4	2.2	7.5	2.03
Dust1 (size range)	2.6	0.65	2.0
Dust2	2.6	1.4	2.0
Dust3	2.6	2.4	2.0
Dust4	2.6	4.5	2.0
Dust5	2.6	8.0	2.0

Hydrophilic aerosol particle size increases as the ambient humidity increases

After the size distribution and refractive index are computed, we apply Mie code to compute mass extinction coefficient ($m^2\ g^{-1}$), single-scattering albedo, and phase function. Finally, AOD for each aerosol type i at the jth atmospheric layer for a wavelength λ can be calculated as

$$\tau_{ij}(\lambda) = \mathrm{ext}(\lambda, i, r_{\mathrm{eff}}) \times c_{ij}, \tag{22}$$

where c_{ij} is aerosol column mass in g/m^2 for an aerosol type i and an atmospheric layer j. It can be seen from Eq. (22) that the AOD depends linearly on the layer column aerosol mass. The column total AOD is a sum of Eq. (22) over all aerosol types and atmospheric layers.

4.6 Clouds

Cloud particle size changes from micrometer to centimeter. The particle size affects emission and scattering of electromagnetic waves from ultraviolet to microwave. In the CRTM, we deal with six cloud types: liquid, ice, hail, graupel, and snow corresponding to the densities in g cm^{-3} of 1.0, 0.9, 0.9, 0.4, and 0.1, respectively. Liquid clouds (water and rain) are assumed to be spherical water droplet and follow the size distribution given by Hansen and Travis (1974):

$$n(r) = r^{(1-3b)/b} e^{-r/ab}, \tag{23}$$

where r is the radius of the water cloud particle, a is the effective radii, and b is the effective variance. The single-scattering properties of water droplets, including the asymmetry factor, single-scattering albedo, extinction efficiency, and scattering phase function, are computed from the Lorenz–Mie program developed by Wiscombe (Wiscombe 1980). Nonspherical particles are used for solid clouds: ice, snow, graupel, and hail. The CRTM adopted the lookup table of optical properties at the microwave range from Liu (2008) and the lookup table given by Yang et al. (2005) for ultraviolet, visible, and infrared spectrum. Table 5 lists cloud size parameters that are used for calculation size distributions of spherical particles.

4.7 Surface Reflectivity and Emissivity Models

The surface emissivity/reflectance models are divided into water, land, ice, and snow gross types for visible, infrared, and microwave sensors. Each gross type is further divided into subtypes, for example, new and old snows. For infrared and visible spectrum, the ASTER spectral library (Baldridge et al. 2009) data is applied for land infrared emissivity and the surface is assumed as Lambertian and the emissivity equals one minus reflectivity (Vogel et al. 2011). We also developed

Table 5 Cloud size parameters used for calculating cloud size distributions

Type	Density g cm^{-3}	MW r_{min} = 0.1 μm, r_{min} = 8000 μm			IR/VIS/UV r_{min} = 0.1 μm, r_{min} = 500 μm		
		Variance of size	Sphere	Nonsphere	Variance of size	Sphere	Nonsphere
Liquid	1.0	0.12 (stratocumulus)	Yes		0.12 (stratocumulus)	Yes	
Rain	1.0	1/3 (Marshall–Palmer) (Liu and Boukabara 2014)	Yes		1/3 (Marshall–Palmer)	Yes	
Ice	0.916	0.18 Nov. 1 cirrus	Yes		0.18 Nov. 1 cirrus	Yes	
Snow	0.1	1/3 (Marshall–Palmer)		Yes	1/3 (Marshall–Palmer)	Yes	
Graupel	0.4	1/3 (Marshall–Palmer)		Yes	1/3 (Marshall–Palmer)	Yes	
Hail	0.916	1/3 (Marshall–Palmer)		Yes	0.1		Yes

utility codes for users to use emissivity atlas that depends on month and latitude and longitude. The infrared water emissivity is based on Wu–Smith model (Van Delst and Wu 2000). The infrared water bidirectional reflectance distribution function (BRDF) model is used for the CRTM direct reflectance to compute reflected solar radiation. The solar radiation over sunglint area may contribute 30 K to infrared channel at 3.7 μm. Microwave water emissivity can be calculated from surface temperature, wind vector, and salinity (Liu et al. 2011). For a calm water surface, surface emissivity is well described by Fresnel equations, as a function of temperature and salinity for a given frequency and a zenith angle. Surface wind generates surface roughness and foam that affect surface emissivity. Both a physical land surface emissivity (Weng et al. 2001) and atlas technique are available in the CRTM. In addition to the physical and atlas emissivity models, surface emissivity over snow and ice surface can be estimated from satellite measured brightness temperature inside the CRTM. The purpose of this model is to use satellite measurements at a few window channels to estimate snow or ice emissivity at those frequencies first, then predict emissivity of sounding channels for those that are also affected by surface emission (Yan et al. 2008).

5 Aerosol Models

A chemical transport model (CTM) is a numerical model that simulates the entire cycle for the species of interest, considering emissions, chemical production and loss, transport and mixing, as well as wet and dry removal. CTMs can be classified accordingly to their species of interest (e.g., aerosol models and dust models), their methodology (e.g., Eulerian models and Lagrangian models), their model characteristics (global models and regional models), as well as their integration strategy

(e.g., offline models whose chemistry is run after the meteorological simulation is done and online models that allow interactive coupling of chemistry and meteorological processes).

This section summarizes the characteristics of two aerosol prediction models that are run in an operational or pseudo-operational manner at NOAA National Centers for Environmental Prediction (NCEP). This chapter is not intended to provide a comprehensive summary of aerosol prediction models running at operational, quasi-operational, or research mode at various operational centers and research institutes around the world. The descriptions of the two aerosol models aim to put CRTM's air quality applications into the contexts.

5.1 Goddard Chemistry Aerosol Radiation and Transport Model (GOCART)

Funded mainly by NASA Earth Science programs, the Goddard Chemistry Aerosol Radiation and Transport model (GOCART) was developed to simulate atmospheric aerosols (including sulfate, black carbon (BC), organic carbon (OC), dust, and sea salt), and sulfur gases (Colarco et al. 2010; Chin et al. 2000, 2003, 2004, 2007, 2009; Ginoux et al. 2001, 2004; Janjic 2003; Bian et al. 2010; Kim et al. 2013).

Aerosol species are assumed to be external mixtures. Total mass of sulfate and hydrophobic and hydrophilic modes of carbonaceous aerosols are tracked, while for dust and sea salt the particle size distribution is explicitly resolved across five noninteracting size bins for each. Both dust and sea salt have dynamic (wind speed dependent) emission functions, while sulfate and carbonaceous species use emission inventory from fossil fuel combustion, biomass burning, and biofuel consumption, with additional biogenic sources of organic carbon. Sulfate has additional chemical production from oxidation of SO_2 and DMS, and a database of volcanic SO_2 emissions and injection heights is used. Aerosol sinks include wet removal (scavenging and rainout) and dry deposition (gravitational sedimentation and surface uptake).

Originally, GOCART was developed as an offline CTM, driven by assimilated meteorological fields from the Goddard Earth Observing System Data Assimilation System [GEOS DAS, e.g., Chin et al. (2002)]. As part of the GEOS Version 4 (GEOS-4) atmospheric model development at NASA Global Modeling and Assimilation Office (GMAO), an online version of GOCART has been developed (Colarco et al. 2010) and was later implemented into NCEP's global model (Lu et al. 2013). When running within versions 4 and 5 of GEOS (GEOS-4/5), the GOCART module provides aerosol processes such as emissions, sedimentation, dry and wet depositions (Fig. 3). Advection, turbulent, and convective transport is outside the GOCART module, being instead provided by the host atmospheric model. Unlike offline CTM, this online aerosol module accurately utilizes winds,

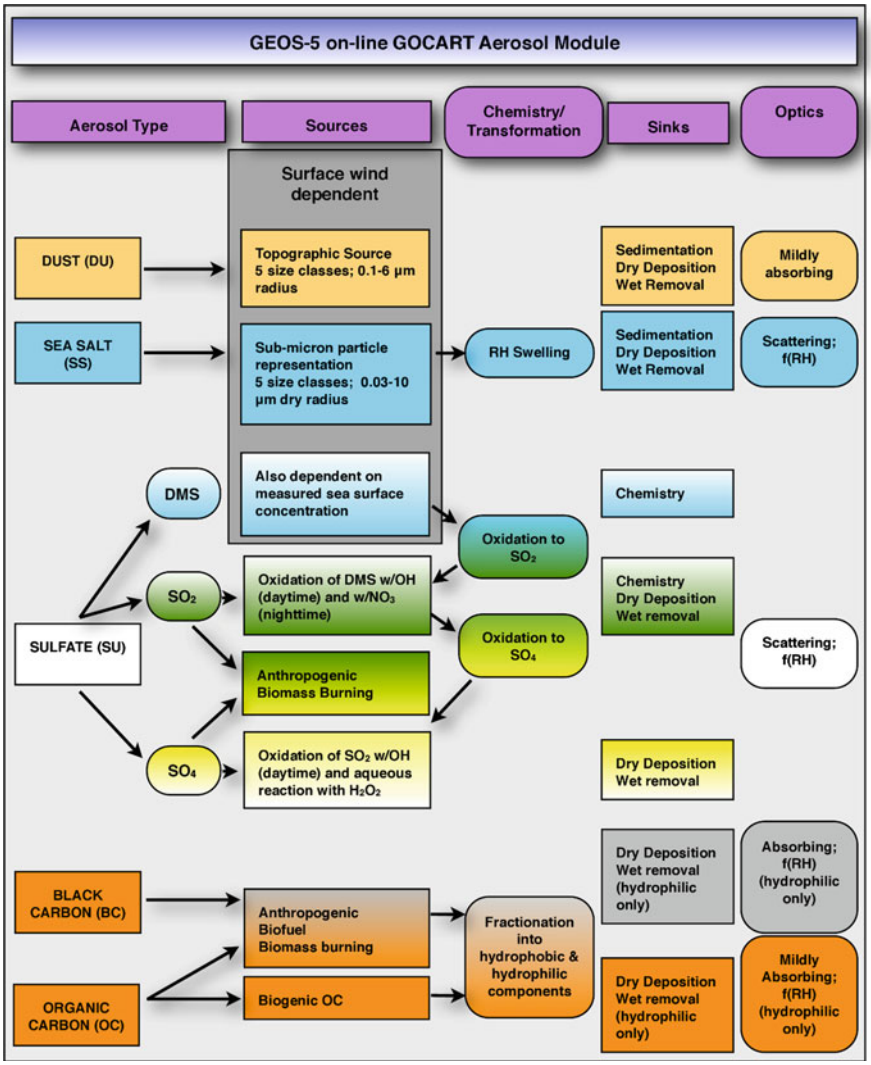


Fig. 3 Schematic summary of the GOCART aerosol modules as adapted and being implemented in GEOS-4/5 at GMAO and NEMS GFS at NCEP

convective mass flux, and eddy diffusivity valid at each time step, without the need for temporal or spatial interpolation of any kind.

Recent research and development efforts have further enhanced GOCART modeling capabilities. The transition to online modeling mentioned above is an example. In addition, the GOCART module now has the option to ingest daily biomass burning emissions from the Quick Fire Emission Dataset (Darmenov and da Silva 2013). QFED emissions are based on fire radiative power retrievals from

MODIS (Moderate Resolution Imaging Spectroradiometer, on board Aqua and Terra satellites). The inclusion of such observation-based, time-dependent emissions is important for model to capture the large temporal–spatial variation of biomass burning emissions.

The GOCART module is currently used in the GEOS-5 Near-Real-Time (NRT) aerosol forecasting system as well as the Modern Era Retrospective analysis for Research and Applications Aerosol Reanalysis (Buchard et al. 2014). At NCEP, the GOCART module is coupled with NCEP's meteorological model for global dust forecasts (described in Sect. 5.3) and an upgrade for multispecies forecast (including dust, sea salt, sulfate, and carbonaceous aerosols) is planned in 2015.

5.2 *The Community Multi-scale Air Quality (CMAQ) Modeling System*

The U.S. Environmental Protection Agency (EPA) has established National Ambient Air Quality Standards (NAAQS), requiring the development of effective emissions control strategies for pollutants such as ozone, particulate matter (PM), and nitrogen species. To meet the emission regulation needs, EPA developed the Community Multi-scale Air Quality (CMAQ) system to develop emission control strategies.

CMAQ is a state-of-the-science "one-atmosphere" system that treats major atmospheric and land processes (e.g., advection, diffusion, nucleation, coagulation, wet and dry deposition, gas phase chemistry, gas–particle mass transfer, and aqueous phase chemistry) for a range of species (e.g., anthropogenic and biogenic, primary and secondary, gaseous and particulate) in a comprehensive framework. The science, model engineering concepts, and infrastructure design have been well documented (Byun et al. 1995a, b, 1996, 1998; Ching et al. 1995; Coats et al. 1995; Byun and Schere 2006).

The CMAQ system framework is comprised of three major modeling systems: emission projection and processing system, the meteorological modeling system, and the CTM. The CMAQ CTM is used to simulate multiple pollutants with emissions and meteorological input data. The science options for CMAQ include the gas phase mechanism, PM module, a set of chemical solvers, various options for vertical and horizontal advection, and photolysis rate calculations.

In response to congressional mandate, NOAA and EPA formed a partnership to transfer scientific advances in air quality modeling at EPA into NCEP's operational model suite. The real-time air quality forecasting system based on CMAQ will be discussed in Sect. 5.4.

5.3 Global Modeling: GOCART in NEMS GFS Aerosol Component (NGAC)

The online version of GOCART module developed for GEOS-4/5 is fairly independent of the host meteorological model, encapsulating the basic aerosol production and loss functionality. It has been incorporated into NOAA Environmental Modeling System (NEMS) to establish the first interactive global aerosol forecasting system, NEMS GFS Aerosol Component (NGAC), at NCEP (Lu et al. 2010, 2013).

The rationale for developing the global aerosol forecasting capabilities at NOAA includes: (1) to improve weather forecasts and climate predictions by taking into account of aerosol effects on radiation and clouds; (2) to improve assimilation of satellite observations by properly accounting for aerosol effects; (3) to provide aerosol (lateral and upper) boundary conditions for regional air quality predictions; (4) to provide a first step toward aerosol data assimilation and reanalysis; and (5) to produce quality aerosol information that address societal needs and stakeholder requirements, e.g., UV index, air quality, ocean productivity, visibility, and sea surface temperature retrievals.

Effective on September 11, 2012, starting with the 0000 Coordinated Universal Time (UTC) cycle, NCEP begins to run and disseminate data from the NGAC version 1 (NGAC V1) at T126 L64 resolution ($\sim$100 km). It provides 5-day dust forecasts, once per day for the 0000 UTC cycle. Dust initial conditions are taken from the 24-h NGAC forecast from previous day while meteorological initial conditions are downscaled from high-resolution Global Data Assimilation System (GDAS) analysis.

The NGAC V1 output is available in GRIdded Binary edition 2 (GRIB2) format on $1 \times 1^{\circ}$ output grid, with 3-hourly output from 00 to 120 h. Primary output fields are global three-dimensional dust mixing ratios for five particle sizes with effective radius at 1, 1.8, 3, 6, and 10 μm. Two-dimensional aerosol diagnosis products, such as aerosol optical depth (AOD) and surface mass concentration, are also available. The NGAC digital products can be accessed from NOMADS at http://nomads.ncep.noaa.gov/ and NCEP's ftp server at ftp://ncep.noaa.gov/pub/data/nccf/com/ngac.

Figure 4 presents the evaluation of NGAC dust distributions. NGAC dust AOD has been compared to dust AOD from GEOS-5 and total AOD from MODIS onboard Terra. The source regions over the Sahara and Sahel are clearly shown as well as the patterns of long-range dust transport. Trade winds steer African dust westward across the Atlantic ocean, covering vast areas of the North Atlantic and sometimes reaching the Americas (e.g., the Caribbean, southeastern USA, Central America, and Amazon basin). This has implications for air quality, public health, climate, and biogeochemical cycle. For instance, about half of the annual dust supply to the Amazon basin is emitted from a single source in the Sahara, the Bodele depression (Koren et al. 2006).

The International Cooperative for Aerosol Prediction (ICAP), consisting of forecasting center model developers and remote sensing data providers, began

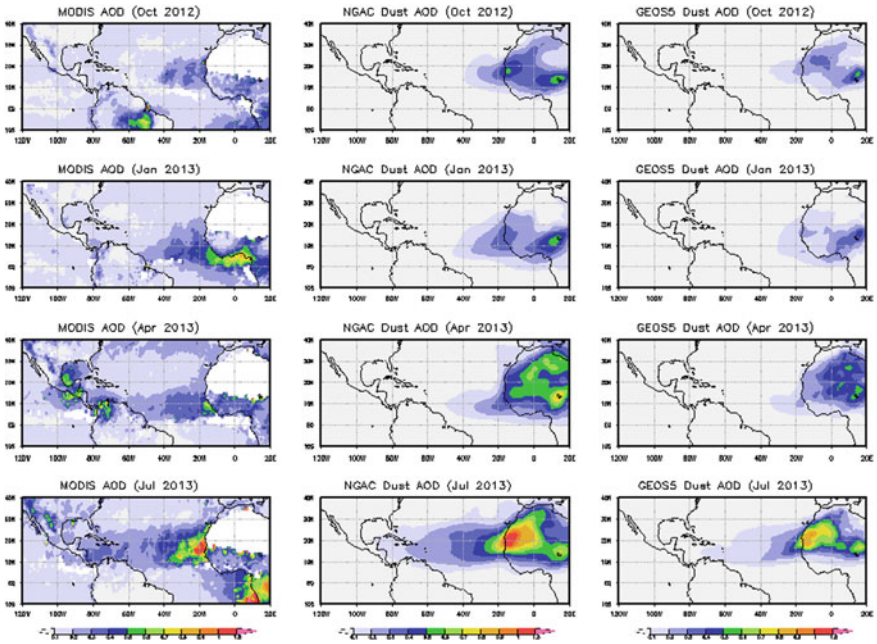


Fig. 4 Comparisons of monthly mean MODIS total AOD (*left*), NGAC dust AOD (*middle*), and GEOS5 dust AOD (*right*) at 550 nm for 2012/10, 2013/01, 2013/04, and 2013/07 periods

meeting in April 2010 to discuss issues relevant to the operational aerosol forecasting (Benedetti et al. 2011; Reid et al. 2011). Consensus ICAP multi-model ensembles [ICAP-MME, (Sessions et al. 2015)] become pseudo-operational since 2014, using four complete aerosol forecast models from GMAO, European Centre for Medium-Range Weather Forecasts (ECMWF), Naval Research Laboratory (NRL), and Japan Meteorological Agency (JMA) as well as three dust-only models from NCEP (i.e., NGAC), U.K. Met Office (UKMO), and Barcelona Supercomputing Center (BSC). Figure 5 shows the dust AOD from ICAP-MME and NGAC, valid at 12 UTC onAugust 1, 2013. Spatial pattern of dust loading from NGAC is consistent with the ICAP-MME, with elevated dust located in the Sahara, the Arabian Peninsula, and Asia.

While NGAC has the capability to predict dust, sulfate, sea salt, and carbonaceous aerosols, the initial operation implementation only provides global dust forecasts. Multispecies aerosol predictions by NGAC require real-time estimates of aerosol precursor emissions. NOAA-NASA has developed global near-real-time (NRT) smoke emissions, blended from NOAA's Global Biomass Burning Emission Product from a constellation of geostationary satellites [GBBEP, (Zhang et al. 2011)] and NASA's Quick Fire Emissions Data from polar orbiting sensor MODIS (QFED). The NRT smoke emissions dataset is targeted to be operational in May 2015, and will be used in the planned NGAC upgrade (multispecies forecasts) slated for operation implementation in late 2015.

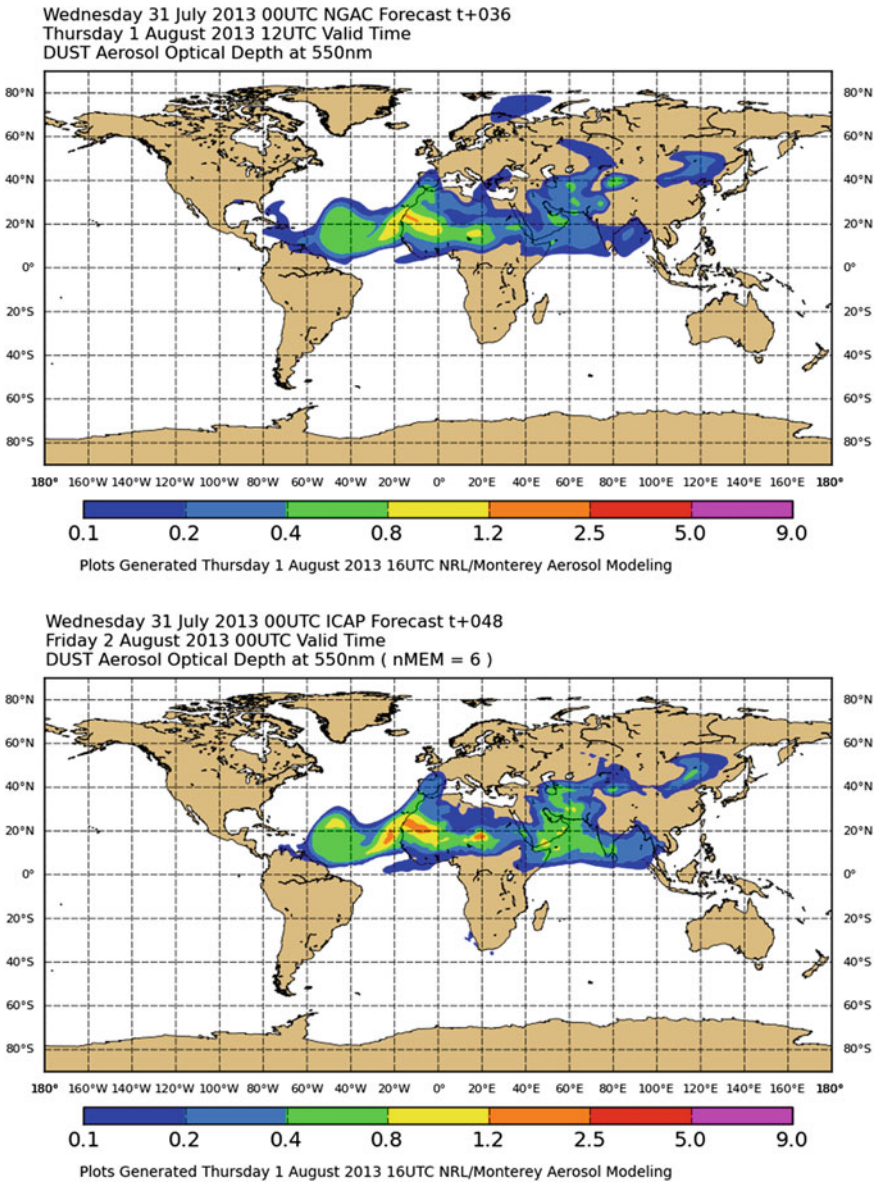


Fig. 5 Dust AOD valid at 12UTC on August 1, 2013 for ICAP multi-model ensemble (*top*) and NGAC (*bottom*). The ensemble is based on six members, including NCEP NGAC, GMAO GEOS-5, ECMWF MACC (Monitoring Atmospheric Composition and Climate), NRL NAAPS (Navy Aerosol Analysis and Prediction System), JMA MASINGAR (Model of Aerosol Species in the Global Atmosphere), and BSC NMMB-CTM. These figures are produced by the Naval Research Laboratory

5.4 Regional Modeling: CMAQ in National Air Quality Forecasting Capability (NAQFC)

As directed by U.S. Congress, NOAA developed a National Air Quality Forecasting (AQF) System (Davidson et al. 2004) through a partnership with the U.S. EPA. The AQF system is an offline coupled atmospheric chemical forecasting system using EPA CMAQ modeling system (Ching et al. 1995), driven by meteorological forecasts from NCEP North American Mesoscale (NAM) system. The real-time U.S. National Air Quality Forecast Capability (NAQFC), based on the NAM-CMAQ system (Otte et al. 2005; Eder et al. 2010; Chai et al. 2013), run at 12 km horizontal resolution out to 48 h twice per day. The initial deployment of NAQFC covered the northeastern U.S. domain since September 2004. It was expanded to cover the eastern U.S. in August 2005 and is now extended to cover the entire continental U.S., Alaska, and Hawaii.

The NAQFC provides operational predictions of ozone from anthropogenic and natural sources since 2004. It will be upgraded to provide PM predictions in 2015. An alternative quick response model for predicting wildfire smoke and dust storms, both of which have highly variable intermittent sources, are based on the Hybrid Single Particle Lagrangian Integrated Trajectory model [HYSPLIT, (Draxler et al. 2010)] and are beyond the scope of this chapter. Descriptions of HYSPLIT smoke and dust forecast systems can be found in Eder et al. (2010) and Rolph et al. (2009), respectively.

The Carbon Bond (CB05) with 51 species and 156 reaction gas phase chemical mechanism is used operationally at NCEP (Sarwar et al. 2008). The modeling of aerosols in CMAQ is based on the PM module (AERO-IV) described in Binkowski and Roselle (2003). Secondary Organic Aerosol formation is included and described in Colarco et al. (2010). The fine PM with size is equal to or less than 2.5 μm in diameter (PM2.5) can be predicted, speciated to nitrate, sulfate, ammonium, organics, and elemental carbon.

The Aerosol Version 4 (AERO-IV) particle model is included in the NAQFC operational configuration for aerosol chemistry. Aerosol size distributions are represented by lognormal distributions of φ, geometric diameter of the particles: Aitken ($\varphi < 0.1$ μm), accumulation ($0.1 < \varphi < 2.5$ μm), and coarse ($2.5 < \varphi < 10$ μm). New particle formation from gas conversion and nucleation are included as well as heterogeneous hydrolysis reaction of N_2O_5 (key linkage between gas and aerosol phase reactions). Predictions of organic nitrate were overpredicted in earlier configurations and influenced ozone production. As of 2015, organic nitrate is photolyzed and removed quicker by shortening lifetime by a factor of 10 based on the findings of Anderson et al. (2014) and Yang et al. (2014). Fugitive dust emissions are included and are now modulated whenever there is ice/snow to suppress emissions. Real-time wildfire smoke emissions are incorporated using the NESIDS Hazard Mapping System "observed" wildfire near-real-time smoke emissions and the U.S. Forest Service BlueSky smoke emissions system.

The emission dataset is built on top of the 2011 National Air Quality Forecasting Capability (NAQFC) baseline data, but with four major updates: (a) new point source emissions with updated emission measurements and energy projections; (b) new mobile source emissions updated to 2012 based on trends from the U.S. EPA surface monitoring network corroborated with satellite trends for the same constituents (Pan et al. 2014); (c) new off-road emissions projected to 2012; and (d) updated Canadian emission sectors from Environment Canada (EC) 2012 emission inventories and the Mexican 2010 National Inventory for Mexico. The NOAA Hazard Mapping System (HMS) is used to detect wildfires over the nation. The HMS product is then used to drive the U.S. Forest Service BlueSky wildfire emissions system. Finally, a Kalman filter-based bias correction scheme (Djalalova et al. 2015) has provided improved PM prediction skill that should result in improved guidance for State environmental agencies responsible for air quality alerts.

The NWS Air Quality Forecast Guidance, updated twice daily, is available at the NWS National Digital Graphical Database (NDGD) decision support system website (http://airquality.weather.gov). U.S. State and local environmental agencies use this product to issue air quality forecasts and AQI predictions in their jurisdictions. To provide the public with easy access to national air quality information, the EPA's AIRNow site (http://www.airnow.gov) was developed. This site provides official AQ point forecasts, issued by state and local AQ forecasters, and real-time AQI conditions.

GEOS-Chem monthly climatological gas and particle species are used at the NAM-CMAQ boundaries. The efforts to use aerosol dynamic lateral boundary conditions (LBCs) from NGAC are under way. An example on using NGAC dust information to improve air quality forecasts is presented here. Two CMAQ runs are conducted for the July 2010 period. The baseline run uses static LBCs and the experimental run uses dynamic LBCs from NGAC. Figure 6 shows the observed and modeled surface PM2.5 at two AIRNOW stations in the southeast region. It is found that the incorporation of LBCs from NGAC reduces model biases and improves correlation. Clearly, the inclusion of long-range dust transport via dynamic LBCs leads to significant improvements in CMAQ forecasts during dust intrusion episodes.

6 Satellite Data Assmilation and Air Quality Forecasting

Aerosols play an important role in weather forecasting and climate change. Except for absorbing aerosols, aerosols reflect part of sunlight back to space. Aerosols directly affect forming clouds. The modeling and prediction of aerosols is associated with a large degree of uncertainty due to uncertainties in the emissions, transport, and its interaction with nonlinear physical processes (e.g., radiative effects, cloud and precipitation formation). Ground-based observing networks have been crucial in validating and improving our understanding of aerosol component

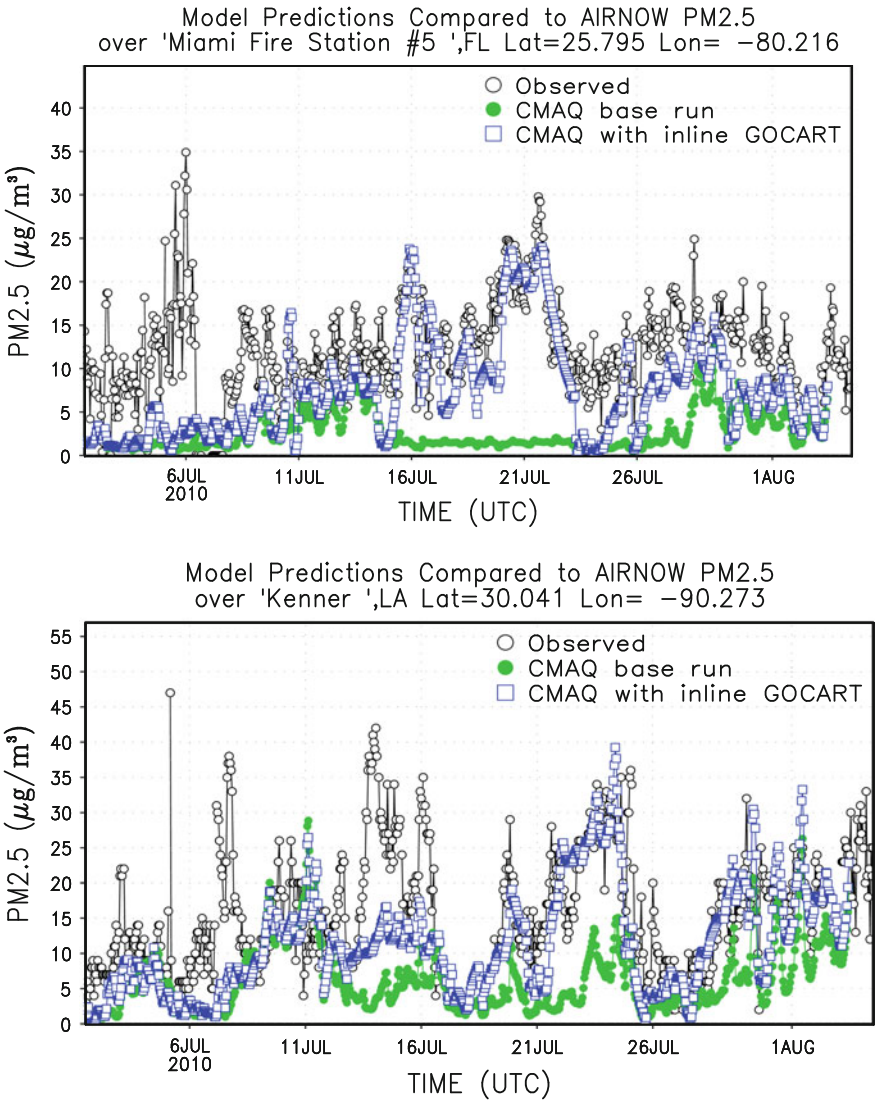


Fig. 6 Time series of PM2.5 from EPA AIRNOW observations (*black dot*), CMAQ baseline run using static LBCs (*green dot*) and CMAQ experimental run using NGAC LBCs (*blue square*) at Miami, FL (*top panel*) and Kenner, LA (*bottom panel*)

of the entire Earth system, but the data cover a very small portion of the Earth. Complemented with ground-based networks, observations from satellite platforms offer a more global view of aerosol distribution. AVHRR (Ignatov et al. 2004), MODIS (Remer et al. 2005), VIIRS (Liu et al. 2014) provide global distribution of aerosol optical depths, which are different from aerosol mass concentration needed

by numerical forecasting model. Data assimilation system can assimilate the aerosol optical depths for aerosol mass concentrations using the CRTM. Data assimilation, an approach combining observations and information from numerical model in a statistically optimal fashion, offers a means to reduce the uncertainties in the estimates of aerosol distribution.

The Grid Statistical Interpolation (GSI) System developed at NCEP (Wu et al. 2002) is used in this study. The core of GSI is to minimize the cost function, which utilizes observations and a priori or forecast information through a 3DVAR method. The data assimilation system by minimizing the cost function finds optimal analysis fields from forecast fields, conventional observations, some retrieval products as observations, and satellite radiances under dynamic constraints following a set of physical laws. The cost function describes the departure of the solution of atmospheric/surface state parameters from background or priori information and the departure of simulations from observations. It may be written as

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x_b})^{\mathbf{T}}\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x_b}) + \frac{1}{2}[H(\mathbf{x}) - \mathbf{y}]^{\mathbf{T}}\mathbf{R}^{-1}[H(\mathbf{x}) - \mathbf{y}] \tag{24}$$

where **B** and **R** are the background and observation error covariance matrices of dimension $n \times n$ and $p \times p$, respectively. They determine the relative weight of the background and observation term contributed to the final analysis. The superscripts "T" and "-1" represent, respectively, the transpose and inverse of the matrix. Here x and y are already assumed to be column vectors and hence $J(x)$ is simply a scalar value.

Most operational NWP centers adopt an incremental implementation (Courtier et al. 1994) of Eq. (24), namely

$$J(\delta\mathbf{x}) = \frac{1}{2}\delta\mathbf{x}^{\mathbf{T}}\mathbf{B}^{-1}\delta\mathbf{x} + \frac{1}{2}(\mathbf{H}\delta\mathbf{x} - \mathbf{d})^{\mathbf{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}) \tag{25}$$

where the analysis is based upon "increment" $\delta\mathbf{x} = \mathbf{x} - \mathbf{x_b}$, and the departure of observation minus background (also referred to as "innovation vector") $\mathbf{d} = \mathbf{y} - H(\mathbf{x_b})$. **H** is the linearized version (or Jacobian, a $p \times n$ matrix) of the nonlinear observation operator H in the vicinity of $\mathbf{x_b}$. The cost function in Eq. (25) is quadratic with respect to $\delta\mathbf{x}$ and thus leads to a faster convergence of a minimization algorithm. More importantly, it allows multiple "outer loops" at lower model resolution for each loop, which is crucial to reduce computational expense in its 4DVAR implementation. At minimum of $J(\delta\mathbf{x})$, the gradient (the first-order partial derivative) of

$J(\delta\mathbf{x})$ with respect to $\delta\mathbf{x}$ needs to be vanished, i.e.,

$$\nabla J_{\delta x} = \mathbf{B}^{-1}\delta\mathbf{x} + \mathbf{H}^{\mathbf{T}}\mathbf{R}^{-1}(\mathbf{H}\delta\mathbf{x} - \mathbf{d}) = 0. \tag{26}$$

In Eq. (5), the background and observation error covariance matrices **B** and **R** are determined offline. In addition to the AOD forward model, the Jacobian of

AOD with respect to aerosol mass concentration is also derived, which is required to calculate the gradient of the cost function (see Eq. 26) in the 3DVAR analysis. In addition to AOD output for those MODIS wavelengths, the CRTM-AOD operator can also calculate AOD for certain wavelengths to validate forecasts with AERONET (500, 675, 870, 1020, and 1640 nm) and CALIOP (532 nm).

The AOD observation operator is based upon the community radiative transfer model (Liu and Weng 2006) developed at the United States Joint Center for Satellite Data Assimilation (JCSDA). CRTM was primarily designed for computing satellite radiances and is used in GSI for directly assimilating radiances from infrared and microwave sensors. However, we extended CRTM to compute MODIS AOD using only aerosol profiles as input. This newly developed CRTM-AOD module was incorporated into the GSI system. Some implementations are ongoing to include WRF-Chem and CMAQ aerosol types.

In CRTM, a lookup table is necessary to store the precalculated aerosol optical parameters such as mass extinction coefficient, single-scattering albedo, and asymmetry factor.

Aerosol data assimilation using AVHRR AOD product into a three-dimensional chemical transport model was introduced by Collins et al. (2001) for studying the INDOEX (Indian Ocean Experiment) aerosols. Liu et al. (2011) used the CRTM in the GSI to assimilate MODIS aerosol optical depths to improve initial fields of aerosol mass concentration.

They adopt the total aerosol mixing ratio as control variable and some assumptions are made in order to partition the total mass into individual species (Benedetti et al. 2009). Its application to a 2-year (2003, 2004) reanalysis using MODIS AOD data showed that the analysis is very skillful in drawing to the observations and in improving the forecasts of AOD.

Experiments with and without the AOD DA have been conducted to evaluate the impact of AOD observations on the aerosol analyses and forecasts of a dust storm period (0000 UTC, 17 ~ 0000 UTC on March 24, 2010) over East Asia. Prior to running the AOD DA experiment, the corresponding background error covariance (BEC) statistics of 14 aerosol species were obtained using traditional "NMC" method, in which the difference of 24-h and 12-h WRF/Chem aerosol forecasts valid at the same time are used for BEC statistics. For the experiment without the AOD DA, aerosol fields are produced by continuous WRF/Chem forecasts driven by GFS meteorological fields and surface emission. AOD DA experiment updates aerosol fields at 0000 and 0600 UTC (day time) when the MODIS has coverage over East Asia. In Fig. 7, the hourly WRF/Chem model output is compared to the independent AOD (1020 nm) observations at two AErosol RObotic NETwork (AERONET) sites (Nanjing of China and Kathmandu of Nepal) for the experimental period. The noDA (green line) experiment severely underestimated AOD. The AOD values from the DA experiment (blue line) agree much more closely with AERONET observations (red line). Maximum dust intensity on March 21, 2010 in Nanjing was well captured by the AOD DA experiment. For Kathmandu site, which was unaffected by the dust storm, the diurnal variation of AOD in the AERONET

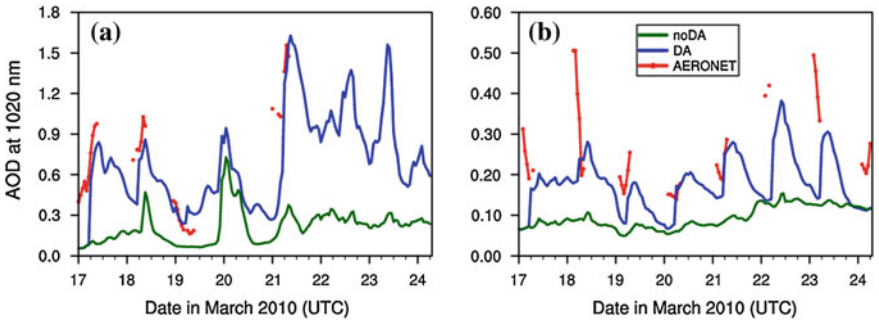


Fig. 7 Hourly time series of AOD at 1020 nm from 0000 UTC 17 to 0600 UTC on March 24, 2010, at Nanjing and Kathmandu AERONET sites. *Red line* denotes the AERONET observations, and *blue* (*green*) curves represent the DA (noDA) experiment, respectively

observations, and replicated by the DA experiment, likely reflects diurnal changes in air quality caused by traffic from morning to evening.

The developed AOD assimilation method can be expanded to assimilate additional aerosol-related observations (e.g., surface measurements of particle matter, multispectral and multi-angle AOD retrievals from different satellite instruments, vertical extinction profiles from ground-based and space-borne Lidar, etc.). Moreover, 3DVAR approach adopted here permits simultaneous assimilation of new aerosol-related data and existing meteorological observations, which are already used in the GSI operation.

Air quality depends on the concentration of trace gases, for example, carbon monoxide is one of the main pollutants. Carbon monoxide is a "combustion pollutant" that comes from incomplete burning carbon materials such as vehicles running, wildfires, and biomass burning emission. Trace gases retrieved from satellite measurements are useful to determine gaseous emission sources. As shown in Fig. 1, CO absorption lines locate between 1150 and 2210 cm^{-1} and observed radiance there can be used to derive CO concentration. Hyperspectral infrared sounding IASI data have been used to derive CO (Pommier et al. 2010). IASI, a Fourier Transform Spectrometer (FTS) records radiance measurements from the Earth's surface and the atmosphere with a high spectral sampling of 0.25 cm^{-1} over a spectral range from 645 to 2760 cm^{-1}, and with a low radiometric noise (0.2–0.35 K at 280 K reference). AIRS data are also used to derive trace gases (Susskind et al. 2003). Based on the AIRS team algorithm (Susskind et al. 2003), the NOAA developed the NOAA Unique CrIS/ATMS Processing System (NUCAPS). The NUCAPS was initially for Suomi NPP CrIS/ATMS sensors and late extends for AIRS and IASI. The description of the AIRS science team algorithm is referred to the paper (Susskind et al. 2003). It needs to be pointed out that the NUCAPS uses the microwave radiative transfer of Rosenkranz (2001) and the Stand-alone AIRS Radiative Transfer Algorithm (SARTA) (Susskind et al. 2003) is used for infrared radiance calculations. Both CRTM and SARTA are channel-based fast radiative transfer models. The NUCAPS algorithm contains two retrieval parts: microwave

(MW) data only, and infrared data with microwave data (IR + MW). The microwave only retrieval (Rosenkranz 2001) can be carried out for all-sky except for strong precipitation cases (Boukabara et al. 2007). The NUCAPS microwave retrieval convergent rate is typically higher than 96 %. Infrared retrieval of atmospheric states is a high nonlinear problem (Larrabee Strow et al. 2003). For retrieval using infrared with microwave data, the first guess plays an important role because of strong nonlinearity of the retrieval problem. We developed two regression algorithms to derive profiles of temperature, water vapor, ozone, surface temperature (Goldberg et al. 2003), and surface infrared emissivity (Zhou et al. 2008). The infrared land surface emissivity may have a large variability (Zhou et al. 2011). The first algorithm is applied for original satellite measurements for all-sky conditions. The regression coefficients are obtained by mapping ECMWF atmospheric state vectors with satellite radiances directly. The second regression algorithm is applied for clear-sky conditions. The NUCAPS algorithm is applied to each field of regard (FOR) of CrIS data where each FOR contains nine (3 × 3) field of view (FOV) in order to collocate with the ATMS measurements. At nadir looking, each CrIS FOV has a footprint of about 14 km in diameter and each CrIS FOR has a spatial resolution of about 45 km. It can be found from Table 6 that the percentage for clear-sky and overcast within FOV is much higher that within FOR. The smaller the footprint, the larger the percentage for clear and overcast scenes. The high percentage is good for clear-sky radiance and cloudy radiance assimilations. The NUCAPS IR + MW retrievals are basically called "clear-sky" retrieval where clear-sky radiances are used. In order to estimate the percentage of clear-sky conditions we collocate the CrIS FOV and FOR with the VIIRS cloud masks (Kopp et al. 2014). The VIIRS data have a good spatial resolution better than 1 knm at nadir (Cao et al. 2013). For one-day global data on May 12, 2014, only 3.3 % is found clear for FORs and 8.6 % is found clear for FOVs. The low percentage of clear-sky radiances caused incomplete information content over some meteorologically very important areas and biasing the model (representative of a clear-sky only issue). To enhance the percentage of IR + MW retrieval, the NUCAPS removes cloud effects from partial cloudy scene, so-called cloud cleared radiance. Cloud cleared radiance technique has been developing and is being used in research communities and satellite retrieval products for decades. The method is applied to remove the cloud effect on radiances for partially cloudy FOV or FOR (Liu et al. 2009). The "clear-sky" radiances availability can be increased from originally 3 % to nearly 70 % for the FOR approach. The NUCAPS algorithm uses stepwise approach to retrieve geophysical parameters sequentially: temperature, water vapor, and ozone. Then the NUCAPS calculates the

Table 6 The data percentage within CrIS field of view (FOV) and field of regard (FOR)

	Clear	1–10	10–20	20–30	30–40	40–50	50–60	60–70	70–80	80–90	90–99	100 % cloud
FOV	8.61	13.92	6.16	5.67	4.52	3.94	3.68	3.80	4.32	8.66	18.28	18.44
FOR	3.30	13.78	7.16	6.78	5.86	5.47	5.33	5.53	6.14	8.52	21.99	10.24

The spatial resolution for FOV and FOR at a nadir looking are about 14 and 45 km, respectively

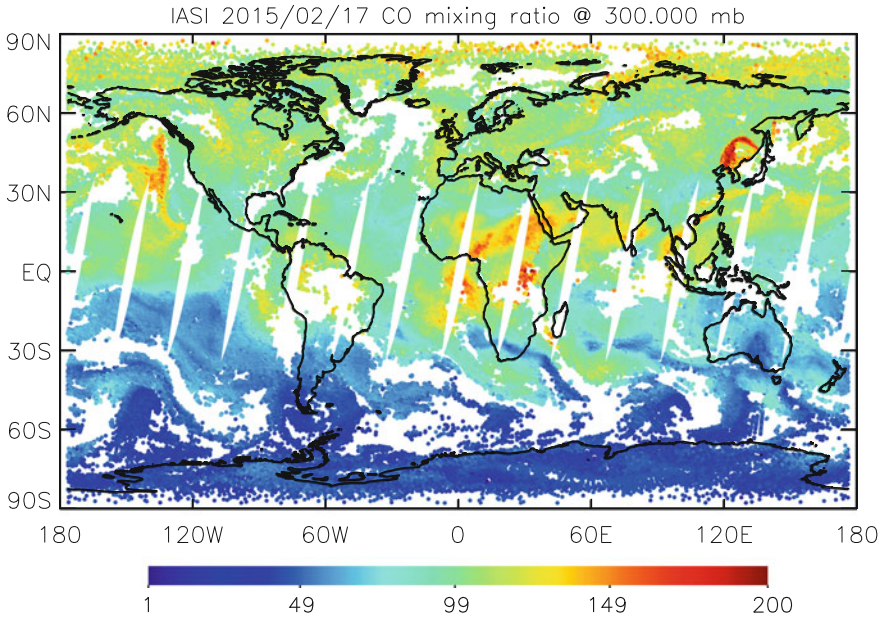


Fig. 8 A global daily carbon monoxide EDR generated by NUCAPS for the IASI measurements

residual between measured and simulated radiances for trace gas sensitive channels (Gambacorta et al. 2014). The residual is used to estimate the concentration of trace gases without changing surface temperature and atmospheric profiles of temperature, water vapor, and ozone.

Figure 8 displayed the global concentration distribution of carbon monoxide from NUCAPS retrieval using the IASI data at ascending and descending orbits, respectively. The IASI data have a good spectral resolution that is particular interesting for retrieving trace gases. One can see high CO concentration located in Southeast Asia and central Africa. The high concentration over Northwest Pacific indicated that CO may transport to northern America.

7 Discussion

This chapter described the Community Radiative Transfer Model (CRTM). The CRTM and Radiative Transfer for TOVS [RTTOV] (Saunders et al. 1999) are the two operational models used for radiance assimilation in supporting of weather forecasting. The CRTM aerosol module has been used in assimilating aerosol optical depth from satellite products and improved air quality forecasting skill (Liu et al. 2011). In this study, we assimilate column aerosol optical depth at the wavelength of 0.55 μm. The vertical distribution of aerosol concentration and

particle size is essential to the aerosol transport. Aerosols of small sizes at high altitude have a large chance to transport a long distance. Since backscattered radiances depend on wavelengths, aerosol optical depths at multiple wavelengths may provide information about the vertical distribution of aerosol concentrations and sizes. It needs to point out that the aerosol definition between the model groups and remote sensing teams may be different. The model group uses aerosol concentrations and particle sizes what dynamic models deal with. The remote sensing team uses optical properties what sensors measure. Even within model community aerosol definition can be different, for example, the size bin of dust is different in the GOCART and the CMAQ models. It is also true within remote sensing community since the retrieval of the aerosol optical depth depends on the assumptions of aerosol type, size, and concentration. Certainly, there is a good correlation between the retrieved aerosol optical depth and the aerosol concentration. To enhance the correlation, the model groups and the remote sensing teams may work together on the common aerosol definition.

The NUCAPS system retrieved trace gases using the IASI data and will generate the trace gaseous EDRs using the CrIS data of a full spectral resolution. The CrIS radiance of a full-spectral resolution is analytically scalable and derivable using fast Fourier transform. The scalable and derivable spectra may provide an additional way to the applications of remote sensing. The spectral resolution is critical to obtain necessary sensitivity of radiances on trace gaseous concentrations. The satellite products provide global distributions of aerosol optical depths and trace gases such as CO and O_3. The data are useful to determine the initial field needed by prediction models. But, the aerosol optical depth represents the column value instead of a vertical distribution. The NUCAPS does provide the vertical distribution of trace gases, but the sensitivity of radiances on trace gases near surface boundary layer is low. From the spectral weighting functions of the IASI and the CrIS, we can see that the most satellite information about trace gases come from the atmospheric layers between 300 and 700 hPa. To improve the retrieval product, good priori information about trace gases is important. Within the data assimilation system, one can integrate broad information from model estimate, ground measures, and satellite radiances. The aerosol vertical distribution may be obtained from aerosol optical depths at multiple wavelengths and at multiple view angles instead of using the optical depth at a single wavelength only. The sensitivity of satellite radiance or reflectance depends on the wavelength. More studies are definitely needed to how to utilize satellite products and how to know the uncertainty of the products for each measurement. The errors estimate will allow users to determine the usage of the products for various applications.

Today's computation capability is large, but it is still not large enough to deal with each gaseous absorption line and strong scatterings of aerosols and clouds. Polarization from clouds and aerosols are not considered because of the shortage of computation capability in an operational system. The errors induced by neglect of the polarization in radiance calculations can be significant [several percents, (Mishchenko et al. 1994)] at UV and visible bands and at microwave frequencies

for some applications. The effect is relatively small at infrared wavelengths. The errors can be as a large of 2 K in microwave brightness temperatures (Liu and Simmer 1996). Surface emissivity models are important in radiative transfer calculations. The uncertainties of the surface emissivity model can be large and degrade the usefulness of image channels that observe surfaces. Multiple scattering calculations can be expansive. With scatterings, radiance calculations may demand several to hundred times more computations than that without scatterings. For operational radiative transfer models, optimizing code and algorithm may not have a big room. We probably need to use advanced technology. GPU may be one of the options that can do designed calculations. It is a known fact that solving linear equations and matrix manipulations including eigenvalue problems are most time-consuming. Therefore, one can focus on those codes for GPU without changing other codes.

Acknowledgments The manuscript contents are solely the opinions of the authors and do not constitute a statement of policy, decision, or position on behalf of NOAA or the U.S. government.

References

Alvarado MJ, Wang C, Prinn RG (2009) Formation of ozone and growth of aerosols in young smoke plumes from biomass burning: 2. Three-dimensional Eulerian studies. J Geophys Res 114:D09307. doi:10.1029/2008JD011186

Anderson DC, Loughner CP, Weinheimer A, Diskin G, Canty TP, Salawitch RJ, Worden H, Fried A, Mikoviny T, Wisthaler A, Dickerson RR (2014) Measured and modeled CO and NOy in DISCOVER-AQ: an evaluation of emissions and chemistry over the eastern US. Atmos Environ 96:78–87

Baldridge AM, Hook SJC, Grove I, Rivera G (2009) The ASTER spectral library version 2.0. Remote Sens Environ 113:711–715. doi:10.1016/j.rse.2008.11.007

Benedetti A et al (2009) Aerosol analysis and forecast in the European centre for medium-range weather forecasts integrated forecast system: 2. Data assimilation. J Geophys Res 114:D13205. doi:10.1029/2008JD011115

Benedetti A, Reid JS, Colarco PR (2011) International cooperative for aerosol prediction (ICAP) workshop on aerosol forecast verification. Bull Am Meteor Soc 92:ES48–ES53. doi:10.1175/BAMS-D-11-00105.1

Bian H, Chin M, Kawa SR, Yu H, Diehl T (2010) Multi-scale carbon monoxide and aerosol correlations from MOPITT and MODIS satellite measurements and GOCART model: implication for their emissions and atmospheric evolutions. J Geophys Res 115:D07302. doi:10.1029/2009JD012781

Binkowski FS, Roselle SJ (2003) Models-3 community multiscale air quality (CMAQ) model aerosol component, 1 model description. J Geophys Res 108:4183. doi:10.1029/2001JD001409

Boukabara S, Weng F, Liu Q (2007) Passive microwave remote sensing of extreme weather events using NOAA-18 AMSUA and MHS. IEEE Geosci Remote Sens 45:2228–2246

Buchard V, da Silva A, Colarco P, Darmenov A, Govindaraju R, Spurr R (2014) Using OMI aerosol index and aerosol absorption optical depth to evaluate the NASA MERRA aerosol reanalysis. Atmos Chem Phys Discuss 14:32177–32231

Byun DW, Schere KL (2006) Review of the governing equations, computational algorithms, and other components of the Models-3 community multiscale air quality (CMAQ) modeling system. Appl Mech Rev 59:51–77

Byun DW, Hanna A, Coats CJ, Hwang D (1995a) Models-3 air quality model prototype science and computational concept development. Transactions of air & waste management association specialty conference on regional photochemical measurement and modeling studies, San Diego, CA, pp 197–212, 8–12 Nov 1993

Byun DW, Coats CJ, Hwang D, Fine S, Odman T, Hanna A, Galluppi KJ (1995b) Prototyping and implementation of multiscale air quality models for high performance computing. Mission earth symposium, Phoenix, AZ, pp 527–532, 9–13 Apr 1993

Byun DW, Dabdub D, Fine S, Hanna AF, Mathur R, Odman MT, Russell A, Segall EJ, Seinfeld JH, Steenkiste P, Young J (1996) Emerging air quality modeling technologies for high performance computing and communication environments. In: Gryning SE, Schiermeier F (eds) Air pollution modeling and its application XI, pp 491–502

Byun DW, Ching JKS, Novak J, Young J (1998) Development and implementation of the EPA models-3 initial operating version: community multi-scale air quality (CMAQ) model: twenty-second NATO/CCMS international technical meeting on air pollution modelling and its application. In: Gryning SE, Chaumerliac N (eds) Air pollution modeling and its application XII. Plenum Publishing Corporation, Berlin, pp 357–368

Cao C, Xiong J, Blonski S, Liu Q, Uprety S, Shao X, Bai Y, Weng F (2013) Suomi NPP VIIRS sensor data record verification, validation, and long-term performance monitoring. J Geophys Res Atmos 118 (2013) doi:10.1002/2013JD020418

Chai T, Kim H-C, Lee P, Tong D, Pan L, Tang Y, Huang J, McQueen J, Tsidulko M, Stajner I (2013) Evaluation of the United States National Air Quality Forecast Capability experimental real-time predictions in 2010 using air quality system ozone and NO_2 measurements. Geosci Model Dev 6:1831–1850. doi:10.5194/gmd-6-1831-2013

Chen Y, Weng F, Han Y, Liu Q (2008) Validation of the community radiative transfer model (CRTM) by using cloudsat data. J Geophys Res 113:D00A03. doi:10.1029/2007JD009561

Chen Y, Han Y, Weng F (2012) Comparison of two transmittance algorithms in the community radiative transfer model: application to AVHRR. J Geophys Res 117:D06206. doi:10.1029/2011JD016656

Chin M, Savoie DL, Huebert BJ, Bandy AR, Thornton DC, Bates TS, Quinn PK, Saltzman ES, De Bruyn WJ (2000) Atmospheric sulfur cycle in the global model GOCART: Comparison with field observations and regional budgets. J Geophys Res 105:24689–24712

Chin M, Ginoux P, Kinne S, Torres O, Holben BN, Duncan BN, Martin RV, Logan JA, Higurashi A, Nakajima T (2002) Tropospheric aerosol optical thickness from the GOCART model and comparisons with satellite and sunphotometer measurements. J Atmos Sci 59:461–483

Chin M, Ginoux P, Lucchesi R, Huebert B, Weber R, Anderson T, Masonis S, Blomquist B, Bandy A, Thornton D (2003) A global aerosol model forecast for the ACE-Asia field experiment. J Geophys Res 108:8654. doi:10.1029/2003JD003642

Chin M, Chu DA, Levy R, Remer LA, Kaufman YJ, Holben BN, Eck T, Ginoux P (2004) Aerosol distribution in the northern hemisphere during ACE-Asia: results from global model, satellite observations, and sunphotometer measurements. J Geophy Res 109:D23S90. doi:10.1029/2004JD004829

Chin M, Diehl T, Ginoux P, Malm W (2007) Intercontinental transport of pollution and dust aerosols: implications for regional air quality. Atmos Chem Phys 7:5501–5517

Chin M, Diehl T, Dubovik O, Eck TF, Holben BN, Sinyuk A, Streets DG (2009) Light absorption by pollution, dust and biomass burning aerosols: a global model study and evaluation with AERONET data. Ann Geophys 27:3439–3464

Ching JKS, Byun DW, Hanna A, Odman T, Mathur R, Jang C, McHenry J, Galluppi K (1995) Design requirements for multiscale air quality models. Mission earth symposium, Phoenix, AZ, pp 532–538, 9–13 Apr 1995

Clough SA, Shephard MW, Mlawer EJ, Delamere JS, Iacono MJ, Cady-Pereira K, Boukabara S, Brown PD (2005) Atmospheric radiative transfer modeling: a summary of the AER codes. J Quant Spectrosc Radiat Transf 91:233–244. doi:10.1016/j.jqsrt.2004.05.058

Coats CJ, Hanna AH, Hwang D, Byun DW (1995) Model engineering concepts for air quality models in an integrated environmental modeling system. Transactions of air & waste management association specialty conference on regional photochemical measurement and modeling studies, San Deigo, CA, pp 213–223, 8–12 Nov 1993

Colarco P, da Silva A, Chin M, Diehl T (2010) Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth. J Geophy Res 115:D14207. doi:10.1029/2009JD012820

Collins WD, Rasch PJ, Eaton BE, Khattatov BV, Lamarque J-F (2001) Simulating aerosols using a chemical transport model with assimilation of satellite aerosol retrievals: Methodology for INDOEX. J Geophys Res 106:7313–7336. doi:10.1029/2000JD900507

Cooke WF, Liousse C, Cachier H, Feichter J (1999) Construction of a 1° x 1° fossil fuel emission data set for carbonaceous aerosol and implementation and radiative impact in the ECHAM4 model. J Geophys Res 104:22137–22162

Courtier P, Thépaut J-N, Hollingsworth J (1994) A strategy for operational implementation of 4D-Var, using an incremental approach. Q J R Meteorol Soc 120:1367–1387. doi:10.1002/qj.49712051912

d'Almeida GA (1991) Atmospheric aerosols, A. Deepak Publishing, Hampton

Darmenov A, da Silva AM (2013) The quick fire emissions dataset (QFED)—documentation of versions 2.1, 2.2 and 2.4, NASA technical report series on global modeling and data assimilation. NASA TM-2013-104606, vol 32, pp 1–183

Davidson PM, Seaman N, Schere K, Wayland RA, Hayes JL, Carey KF (2004) National air quality forecasting capability: first steps toward implementation. In: Proceedings of sixth conference on atmospheric chemistry, American Meteorological Society, Seattle, WA (Paper J2.10)

Ding S, Yang P, Weng F, Liu Q, Han Y, van Delst P, Li J, Baum B (2011) Validation of the community radiative transfer model. J Q S Spectrosc Radiat Transf 112:1050–1064

Djalalova IL, Monache D, Wilczak J (2015) PM2.5 analog forecast and Kalman filter post-processing for the community multiscale air quality (CMAQ) model. Atm Envir 108:76–87

Draxler RR, Ginoux P, Stein AF (2010) An empirically derived emission algorithm for wind blown dust. J Geophys Res 115. doi:10.1029/2009JD013167

Duncan BN, Martin RV, Staudt AC, Yevich R, Logan JA (2003) Interannual and seasonal variability of biomass burning emissions constrained by satellite observations. J Geophys Res 108:4100. doi:10.1029/2002JD002378

Eder B, Kang D, Trivikrama Rao S, Mathur R, Yu S, Otte T, Schere K, Wayland R, Jackson S, Davidson P, McQueen J, Bridgers G (2010) Using national air quality forecast guidance to develop local air quality index forecasts. Bull Am Meteor Soc 91:313–326. doi:10.1175/2009BAMS2734.1

Evans KF, Stephens GL (1991) A new polarized atmospheric radiative transfer model. J Quant Spectrosc Radiat Transfer 46:413–423

Fischer J, Grassl H (1984) Radiative transfer in an atmospheric-ocean system: an azimuthally depedent matrix operator approach. Appl Opt 23:1032–1039

Gambacorta A, Barnet C, Wolf W, King T, Maddy E, Strow L, Xiong X, Nalli N, Goldberg M (2014) An experiment using high resolution NPP CrIS measurements for atmospheric trace gases: carbon monoxide retrievals impact study. IEEE Geosci Remote Sens Lett 11:1639–1643

Ginoux P, Chin M, Tegen I, Prospero J, Holben B, Dubovik O, Lin S-J (2001) Sources and global distributions of dust aerosols simulated with the GOCART model. J Geophys Res 106:20255–20273

Ginoux P, Prospero J, Torres O, Chin M (2004) Long-term simulation of dust distribution with the GOCART model: correlation with the North Atlantic oscillation, environ. Model Softw 19:113–128

Goldberg M, Qu L, McMillin Y, Wolf W, Zhou L, Divakarla M (2003) Airs near-real-time products and algorithms in support of operational weather prediction. IEEE Trans Geosci Remote Sens 41:379–389

Guenther AC, Hewitt N, Erickson D, Fall R, Geron C, Graedel T, Harley P, Graedel L, Lerdau M, McKay WA, Pierce T, Scholes B, Steinbrecher R, Tallamraju R, Taylor J, Zimmerman P (1995) A global model of natural volatile organic compound emissions. J Geophys Res 100:8873–8892

Hale GM, Querry MR (1973) Optical constants of water in the 200-nm to 200-mm wavelength region. Appl Opt 12:555–563

Han Y, van Delst P, Liu Q, Weng F, Yan B, Treadon R, Derber J (2006) Community radiative transfer model (CRTM)—Version 1. NOAA NESDIS Technical Report 122

Han Y, Weng F, Liu Q, van Delst P (2007a) A fast radiative transfer model for SSMIS upper atmosphere sounding channel. J Geophys Res 112:D11121. doi:10.1029/2006JD008208

Han Y, Weng F, Liu Q, van Delst P (2007b) A fast radiative transfer model for SSMIS upper atmosphere sounding channel. J Geophys Res 112:D11121. doi:10.1029/2006JD008208

Hansen JE, Travis LD (1974) Light scattering in planetary atmospheres. Space Sci Rev 16 (1973):527–610

Heidinger AK, Christopher O, Bennartz R, Greenwald T (2006) The successive-order-of-interaction radiative transfer model. Part I: model development. J Appl Meteorol 45:1388–1402

Hess M, Koepke P, Schult I (1998) Optical properties of aerosols and clouds: the software package OPAC. Bull Am Met Soc 79:831–844

Hu YX, Wielicki B, Lin B, Gibson G, Tsay SC, Stamnes K, Wong T (2000) δ-Fit: a fast and accurate treatment of particle scattering phase functions with weighted singular-value decomposition least-squares fitting. JQSRT 65:681–690

Ignatov A, Sapper J, Laszlo I, Nalli N, Kidwell K (2004) Operational aerosol observations (AEROBS) from AVHRR/3 on board NOAA-KLM satellites. J Atmos Oceanic Technol 21 (2004):3–26. doi:10.1175/1520-0426021<0003:OAOAFO>2.0.CO;2

Janjic ZI (2003) A nonhydrostatic model based on a new approach. Meteorol Atmos Phys 82:271–285. doi:10.1007/s00703-001-0587-6

Kim D, Chin M, Bian H, Tan Q, Brown ME, Zheng T, You R, Diehl T, Ginoux P, Kucsera T (2013) The effect of the dynamic surface bareness to dust source function, emission, and distribution. J Geophys Res 118:1–16. doi:10.1029/2012JD017907

Kopp TJ, Thomas W, Heidinger AK, Botambekov D, Frey RA, Hutchison KD, Iisager BD, Brueske K, Reed B (2014) The VIIRS cloud mask: progress in the first year of S-NPP toward a common cloud detection scheme. J Geophys Res Atmos 119:2441–2456. doi:10.1002/2013JD020458

Koren I, Kaufman YJ, Washington R, Todd MC, Rudich Y, Vanderlei Martins J, Resenfeld D (2006) The Bodele depressions: a single spot in the Sahara that provides most of the mineral dust to the Amazon forecast. Environ Res Lett 1:014005 (5pp). doi:10.1088/1748-9326/1/1/014005

Larrabee Strow L, Hannon SE, De Souza-Machado S, Motteler HE, Tobin E (2003) An overview of the AIRS radiative transfer model. IEEE Trans Geosci Remote Sens 41:303–313

Lee P, Liu Y (2014) Preliminary evaluation of a regional atmospheric chemical data assimilation system environmental surveillance. Int J Environ Res Public Health 11:12795–12816

Liang S, Zhong B, Fang H (2006) Improved estimation of aerosol optical depth from MODIS imagery over land surfaces. Remote Sens Environ 104:416–425

Liang X-M, Ignatov A, Kihai Y (2009) Implementation of the community radiative transfer model in advanced clear-sky processor for oceans and validation against nighttime AVHRR radiances. J Geophys Res 114:D06112. doi:10.1029/2008JD010960

Liou KN (2002) An introduction to atmospheric radiation, 2nd edn. Academic Press, San Diego

Liss PS, Merlivat L (1986) Air-sea gas exchange rates: introduction and synthesis. In: Buat-Menard P (ed) The role of air-sea exchange in geochemical cycling. Reidel, Hinghan, MA, pp 113–127

Liu G (2008) A database of microwave single-scattering properties for nonspherical ice particles. Bull Am Meteor Soc. 89:1563–1570. doi:10.1175/2008BAMS2486.1

Liu Q, Boukabara S (2014) Community radiation transfer model (CRTM) applications in supporting the Suomi national polar-orbiting partnership (SNPP) mission validation and verification. Remote Sen Environ 140:744–754

Liu Q, Ruprecht E (1996) A radiative transfer model: matrix operator method. Appl Opt 35:4229–4237

Liu Q, Simmer C (1996) Polarization and intensity in microwave radiative transfer model. Contrib Atmos Phys 69:535–545

Liu Q, Weng F (2006) Advanced doubling-adding method for radiative transfer in planetary atmospheres. J Atmos Sci 63:3459–3465

Liu Q, Weng F (2009) Recent stratospheric temperature observed from satellite measurements. SOLA 5:53–56. doi:10.2151/sola.2009-014

Liu Q, Weng F (2013) Using advanced matrix operator (AMOM) in community radiative transfer. IEEE JSTAR 6:1211–1212. doi:10.1109/JSTARS.2013.2247026

Liu Q, Xiao S (2014) Effects of spectral resolution and signal-to-noise ratio of hyperspectral sensors on retrieving atmospheric parameters. Opt Lett 39:60–63

Liu Z, Vaughan M, Winker D, Kittaka C, Getzewich B, Kuehn R, Omar A, Powell K, Trepte C, Hostetler C (2009a) The CALIPSO lidar cloud and aerosol discrimination: version 2 algorithm and initial assessment of performance. J Atmos Oceanic Technol 26:1198–1213

Liu X, Zhou DK, Larar AM, Smith WL, Schluessel P, Newman SM, Taylor JP, Wu W (2009b) Retrieval of atmospheric profiles and cloud properties from IASI spectra using super-channels. Atmos Chem Phys 9:9121–9142. doi:10.5194/acp-9-9121-2009

Liu Z, Liu Q, Lin HC, Schwartz CS, Lee YH, Wang T (2011a) Three-dimensional variational assimilation of MODIS aerosol optical depth: implementation and application to a dust storm over East Asia. J Geophys Res 116:D23206. doi:10.1029/2011JD016159

Liu Q, Weng F, English S (2011b) An improved fast microwave water emissivity model. IEEE TGRS 49:1238–1250

Liu Q, Li C, Xue Y (2013) Sensor-based clear and cloud radiance calculations in the community radiative transfer model. Appl Opt 52:4981–4990

Liu H, Remer LA, Huang J, Huang H-C, Kondragunta S, Laszlo I, Oo M, Jackson JM (2014) Preliminary evaluation of S-NPP VIIRS aerosol optical thickness. J Geophys Res Atmos 119 (2014):3942–3962. doi:10.1002/2013JD020360

Lu S, Huang H-C, Hou Y-T, Tang Y, McQueen J, da Silva A, Chin M, Joseph E, Stockwell W (2010) Development of NCEP global aerosol forecasting system: an overview and its application for improving weather and air quality forecasts. In: NATO science for peace and security series: air pollution modelling and its application vol XX, pp 451–454. doi:10.1007/978-90-481-3812-8, 2010

Lu S, da Silva A, Chin M, Wang J, Moorthi S, Juang H, Chuang HY, Tang Y, Jones L, Iredell M, McQueen J (2013) The NEMS GFS aerosol component: NCEP's global aerosol forecast system, NCEP Office Note 472. Available at: http://www.lib.ncep.noaa.gov/ncepofficenotes/files/on472.pdf, Washington D.C., 26 pp

Lu S, Iredell M, Wang J, Moorthi S, McQueen J, Chuang H-Y, Hou Y-T, Juang H, Yang W, da Silva A, Chin M (2013) The NEMS GFS aerosol component: NCEP's global aerosol forecast system, NCEP Office Note 472, 26 pp. Available at: http://www.lib.ncep.noaa.gov/ncepofficenotes/files/on472.pdf, Washington D.C.

McMillin LM, Crone JJ, Goldberg MD, Kleespies TJ (1995) Atmospheric transmittance of an absorbing gas. 4. OPTRAN: a computationally fast and accurate transmittance model for absorbing gases with fixed and variable mixing ratios at variable viewing angles. Appl Opt 34:6269–6274

Mishchenko MI, Lacis AA, Travis LD (1994) Errors induced by the neglect of polarization in radiance calculations for Rayleigh-scattering atmospheres. J Quant Spectrosc Radiat Transfer 51:491–510

Mishchenko MI, Travis LD, Lacis AA (2006) Multiple scattering of light by particles. University Press, Cambridge

Olivier JG, Bouwman AF, van der Maas CW, Berdowski JJ (1994) Emission database for global atmospheric research (Edgar). Environ Monit Assess 31:93–106. doi:10.1007/BF00547184

Otte TL, Pouliot G, Pleim JE, Young JO, Schere KL, Wong DC, Lee PCS, Tsidulko M, McQueen JT, Davidson P, Mathur R, Chuang HY, DiMego G, Seaman NL (2005) Linking the eta model with the community multiscale air quality (CMAQ) modeling system to build a national air quality forecasting system. Weather Forecast 20:367–384

Pan L, Tong D, Lee P, Kim H-C, Chai T (2014) Assessment of NOx and O_3 forecasting performances in the U.S. national air quality forecasting capability before and after the 2012 major emissions updates. Atmos Envir 95:610–619. doi:10.1016/j.atmosenv.2014.06.020

Plass GN, Kattawar W, Catchings FE (1973) Matrix operator theory of radiative transfer, 1: rayleigh scattering. Appl Opt 12:314–329

Pommier M, Law KS, Clerbaux C, Turquety S, Hurtmans D, Hadji-Lazaro J, Coheur P-F, Schlager H, Ancellet G, Paris J-D, Nédélec P, Diskin GS, Podolske JR, Holloway JS, Bernath P (2010) IASI carbon monoxide validation over the Arctic during POLARCAT spring and summer campaigns. Atmos Chem Phys 10:10655–10678. doi:10.5194/acp-10-10655-2010

Potter P, Ramankutty N, Bennett EM, Donner SD (2010) Characterizing the spatial patterns of global fertilizer application and manure production. Earth Interact 14:1–22. doi:10.1175/2009EI288.1

Reid JS, Benedetti A, Colarco PR, Hansen JA (2011) International operational aerosol observability workshop. Bull Am Meteor Soc 92:ES21–ES24. doi: 10.1175/2010BAMS3183.1

Remer LA et al (2005) The MODIS aerosol algorithm, products, and validation. J Atmos Sci 62:947–973. doi:10.1175/JAS3385.1

Rolph GD, Draxler RR, Stein AF, Taylor A, Ruminski MG, Kondragunta S, Zeng J, Huang H, Manikin G, McQueen JT, Davidson PM (2009) Description and verification of the NOAA smoke forecasting system: the 2007 fire season. Weather Forecast 24:361–378

Rosenkranz PW (2001) Retrieval of temperature and moisture profiles from AMSU-A and AMSU-B measurements. IEEE Trans Geosci Remote Sens 39:2429–2435

Sarwar G, Luecken D, Yarwood G, Whitten G, Carter B (2008) Impact of an updated carbon bond mechanism on air qual-ity using the community multiscale air quality modeling system: preliminary assessment. J Appl Meteorol Clim 47:3–14

Saunders RW, Matricardi M, Brunel P (1999) An improved fast radiative transfer model for assimilation of satellite radiance observations. Quart J R Meteorol Soc 125:1407–1425

Saunders R, Brunel P, von Engeln A, Bormann N, Strow L, Hannon S, Heilliette S, Liu X, Miskolczi F, Han Y, Masiello G, Moncet JL, Uymin G, Sherlock V, Turner DS (2007) A comparison of radiative transfer models for simulating AIRS radiances. J Geophys Res 112: D01S90. doi:10.1029/2006JD007088

Schmetz J, Raschke E (1981) An approximate computation of infrared radiative fluxes in a cloudy atmosphere. Pure appl Geophys 119:248–258

Seinfeld JH, Pandis SN (2006) Atmospheric chemistry and physics—from air pollution to climate change, 2nd edn. Wiley, New York

Sessions WR, Reid JS, Benedetti A, Colarco PR, da Silva A, Lu S, Sekiyama T, Tanaka TY, Baldasano JM, Basart S, Brooks ME, Eck TF, Iredell M, Hansen JA, Jorba OC, Juang H-M, Lynch P, Morcrette J-J, Moorthi S, Mulcahy J, Pradhan Y, Razinger M, Sampson CB, Wang J, Westphal DL (2015) Development towards a global operational aerosol consensus: basic climatological characteristics of the international cooperative for aerosol prediction multi-model ensemble (ICAP-MME). Atmos Chem Phys 15:355–362. doi:10.5194/acp-15-335-2015

Stamnes K, Tsay S-C, Wiscombe W, Jayaweera K (1988) Numerically stable algorithm for discrete ordinate method radiative transfer in multiple scattering and emitting layered media. Appl Opt 27:2502–2529

Susskind J, Barnet CD, Blaisdell J (2003) Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds. IEEE Trans Geosci Remote Sens 41:390–409

Sutton MA, Mason KE, Sheppard LJ, Sverdrup H, Haeuber R, Hicks WK (eds) (2014) Nitrogen deposition, critical loads and biodiversity. Springer, New York

Van de Hulst HC (1963) A new look at multiple scattering, Technical Report. Goddard Institute for Space Studies, NASA, New York

Van Delst P, Wu X (2000) A high resolution infrared sea surface emissivity database for satellite applications.In: Technical proceedings of eleventh international ATOVS study conference, Budapest, Hungary, pp 407–411, 20–26 Sept

Vogel R, Liu Q, Han Y, Weng F (2011) Evaluating a satellite-derived global infrared land surface emissivity data set for use in radiative transfer modeling. J Geophys Res 116:D08105. doi:10.1029/2010JD014679

Weng F, Liu Q (2003) Satellite data assimilation in numerical weather prediction models, part I: forward radiative transfer and jocobian modeling in cloudy atmospheres. J Atmos Sci 60:2633–2646

Weng F, Yan B, Grody NC (2001) A microwave land emissivity model. J Geophys Res 106:20,115–20,123

Wiscombe WJ (1980) Improved Mie scattering algorithms. Appl Opt 19:1505–1509

Wiscombe WJ, The Delta-M Method (1957) Rapid yet accurate radiative flux calculations for strongly asymmetric phase function. J Atmos Sci 34:1408–1422

Wu W-S, Purser RJ, Parrish DF (2002) Three-dimensional variational analysis with spatially inhomogeneous covariances. Mon Weather Rev 130:2905–2916

Wu X, Liu Q, Zeng J, Grotenhuis M, Qian H, Caponi M, Flynn L, Jaross G, Sen B, Buss R et al (2014) Evaluation of the sensor data record from the nadir instruments of the ozone mapping profiler suite (OMPS). J Geophys Res-ATMOS 119:6170–6180

Yan B, Weng F, Meng H (2008) Retrieval of snow surface microwave emissivity from the advanced microwave sounding unit. J Geophys Res 113:D19206. doi:10.1029/2007JD009559

Yang P, Wei HL, Huang HL, Baum BA, Hu YX, Kattawar GW, Mishchenko MI (2005) Fu, Scattering and absorption property database for nonspherical ice particles in the near- through far-infrared spectral region. Appl Opt 44:5512–5523

Yang K, Simon A, Ge CC, Wang J, Dickerson RR (2014) Advancing measurements of tropospheric NO_2 from space: New algorithm and first global results from OMPS. Geophys Res Lett 41(2014):4777–4786. doi:10.1002/2014GL060136

Zhang X, Kondragunta X, Ram J, Schmidt C, Huang H-C (2011) Near-real-time global biomass burning emissions products from geostationary satellite constellation. J Geophys Res 117:D14201. doi:10.1029/2012JD017459

Zhou L, Goldberg M, Barnet C, Cheng Z, Sun F, Wolf W, King T, Liu X, Sun H, Divakarla M (2008) Regression of surface spectral emissivity from hyperspectral instruments. IEEE Trans Geosci Remote Sens 46:328–333

Zhou DK, Larar AM, Liu X, Smith WL, Larrabee Strow L, Yang P, Schlüssel P, Calbet X (2011) Global land surface emissivity retrieved from satellite ultraspectral IR measurements. IEEE Trans Geosci Remote Sens 49:1277–1290. doi:10.1109/TGRS.2010.2051036

Analytical Solution of Radiative Transfer Using Cumulant Expansion

Wei Cai and Min Xu

1 Introduction

Radiative transfer describes light propagation in a turbid medium, where photons suffer multiple scattering by randomly distributed scatterers in the medium. The kinetic equation governing photon propagation is the classic Boltzmann transport equation. The search for an analytical solution of the time-dependent elastic Boltzmann transport equation has lasted for many years. Besides being considered as a classic problem in fundamental research in statistical dynamics, the study of analytical solution of this equation has applications in a broad variety of fields from astrophysics, to geophysics, remote sensing, and medicine. One classic approach to solve this equation is to transform the Boltzmann transport equation into a series of moment equations based on the angular moment expansion and to retain the leading orders (Chandrasekhar 1960; Peraiah 2002; Wehrse and Kalkofen 2006). To the second order, a diffusion equation is derived. One of the main advantages of the moment equations (and related methods such as the diffusion approximation) is that they are manifestly energy conservative. The diffusion approximation admits an analytical solution in a uniform medium of infinite, semi-infinite, slab, and other geometries and has been widely used (Aronson 1995; Cuccia et al. 2009; Xu et al.

W. Cai (✉)
Physics Department, The City College of New York, 136th St and Convent Av., New York, NY 10037, USA
e-mail: weicai593@gmail.com

M. Xu
Physics Department, Fairfield University, 1073 N Benson Rd, Fairfield, CT 06824, USA
e-mail: mxu@fairfield.edu

© Springer-Verlag Berlin Heidelberg 2016

A. Kokhanovsky (ed.), *Light Scattering Reviews, Volume 11*,
Springer Praxis Books, DOI 10.1007/978-3-662-49538-4_3

2001). The diffusion approximation, however, fails at early times when the particle distribution is still highly anisotropic. Numerical approaches, including the Monte Carlo simulation (Patwardhan et al. 2005; Xu 2004; Sawicki et al. 2008), remain the main tools in solving the elastic Boltzmann equation, which are cumbersome tasks.

We present here an analytical solution of the *classic* elastic Boltzmann transport equation in an infinite uniform medium, with the photon's velocity $c\mathbf{s}$, where $\mathbf{s}$ is a unit vector of direction, and c is the (constant) speed in the medium (Cai et al. 2000a, b, 2002, 2003, 2005, 2006; Xu et al. 2001, 2002, 2004; Liemert and Kienle 2012; Cai and Gayen 2010). We assume that the phase function depends only on the scattering angle. Under this assumption, we can handle an arbitrary phase function for obtaining the photon distribution (specific intensity), $I(\mathrm{r}, \mathbf{s}, t)$, as a function of time t, position $\mathbf{r}$, and direction $\mathbf{s}$, and the particle density distribution $N(\mathrm{r}, t)$.

Our approach is outlined as follows. First, the exact expression of the total angular distribution, $F(\mathbf{s}, \mathbf{s}_0, t)$, as a function of time in an infinite uniform medium is derived where $\mathbf{s}_0$ is the incident direction of the beam at time $t = 0$. Based on this angular distribution, we then derive exact spatial cumulants of $I(\mathbf{r}, \mathbf{s}, t)$ up to an arbitrary order at any angle and time (Cai et al. 2000a, 2005; Xu et al. 2002). A cutoff at the second order of cumulants, $I(\mathbf{r}, \mathbf{s}, t)$ and $N(\mathbf{r}, t)$, can be expressed by Gaussian distributions (Baum et al. 2000; Cai et al. 2000a, 2003), which have exact first cumulant (the position of the center of the distribution) and exact second cumulant (the half-width of the spread of the distribution). The solution has been extended to the case of polarized photon distribution (Cai et al. 2000b, 2006). By use of a perturbative method, the distribution in a weak heterogeneous scattering medium can be computed (Cai et al. 2003; Xu et al. 2001, 2002).

This chapter is organized as follows: In Sect. 2, expressions of spatial cumulants of the distribution up to an arbitrary order are derived. In Sect. 3, the explicit Gaussian forms of the distribution (up to the second-order cumulant) are presented, and numerical results of the solutions are shown. In Sect. 4, the cumulant solution for polarized light transfer is presented. Numerical solutions for backscattering of circular polarized light are presented, comparing with experimental results. In Sect. 5, applications of the cumulant solution of radiative transfer in areas of biology tomography using early time gates technology (Niedre et al. 2008), and of remote sensing for determining size and density of water drop in cloud (Cai and Gayen 2010) are presented. This chapter is finally summarized in Sect. 6.

2 Derivation of Cumulants to an Arbitrary Order

The elastic Boltzmann kinetic equation of light for the distribution function $I(\mathrm{r}, \mathbf{s}, t)$ as a function of time t, position $\mathbf{r}$, and direction $\mathbf{s}$, in an infinite uniform medium, from a point pulse light source $\delta(\mathbf{r} - \mathbf{r_0})\delta(\mathbf{s} - \mathbf{s_0})\delta(t - 0)$ is given by

$$\partial I(\mathbf{r},\mathbf{s},t)/\partial t + c\mathbf{s}\cdot\nabla_{\mathrm{r}} I(\mathbf{r},\mathbf{s},t) + \mu_a I(\mathbf{r},\mathbf{s},t) = \mu_s \int P(\mathbf{s},\mathbf{s}')[I(\mathbf{r},\mathbf{s}',t) - I(\mathbf{r},\mathbf{s},t)]\mathrm{d}\mathbf{s}' + \delta(\mathbf{r}-\mathbf{r_0})\delta(\mathbf{s}-\mathbf{s_0})\delta(t-0), \tag{2.1}$$

where c is the speed of light, μ_s is the scattering rate, μ_a is the absorption rate, and $P(\mathbf{s}, \mathbf{s'})$ is the phase function, normalized to $\int P(s,s')\mathrm{d}s' = 1$.

When the phase function depends only on the scattering angle in an isotropic medium, we can expand the phase function in Legendre polynomials:

$$P(\mathbf{s},\mathbf{s}') = \frac{1}{4\pi}\sum_l a_l P_l[\cos(\mathbf{s}\cdot\mathbf{s}')]. \tag{2.2}$$

where a_l are coefficients determined by the phase function. A difficulty in solving Eq. (2.1) originates from the term

$$c\mathbf{s}\cdot\nabla_r I(\mathbf{r},\mathbf{s},t),$$

which couples components of $I(r, \mathbf{s}, t)$—spherical harmonics of different orders with each other. We first study the dynamics of the distribution in the direction space, $F(\mathbf{s}, \mathbf{s}_0, t) = \int I(\mathbf{r},\mathbf{s},t)\mathrm{d}\mathbf{r}$, on a spherical surface of unit radius. The kinetic equation for $F(\mathbf{s}, \mathbf{s}_0, t)$ can be obtained by integrating Eq. (2.1) over the whole spatial space, $\mathbf{r}$. The spatial independence of μ_S, μ_a, and $P(\mathbf{s}, \mathbf{s'})$ ensures translation invariance. Thus the integral of Eq. (2.1) obeys

$$\partial F(s,s_0,t)/\partial t + \mu_a F(s,s_0,t) + \mu_s\left[F(s,s_0,t) - \int P(s,s')F(s',s_0,t)\mathrm{d}s'\right] = \delta(s-s_0)\delta(t-0) \tag{2.3}$$

since the integral of the gradient term over all space vanishes. In contrast to Eq. (2.1), upon expanding $F(\mathbf{s}, \mathbf{s'}, t)$ in spherical harmonics, its components do not couple with each other. Therefore, it is easy to obtain the exact solution of Eq. (2.3):

$$F(s,s_0,t) = \exp(-\mu_a t)\sum_l \frac{2l+1}{4\pi}\exp(-g_l t)P_l[\cos(s\cdot s_0)], \tag{2.4}$$

where

$$g_l = \mu_s[1 - a_l/(2l+1)]. \tag{2.5}$$

Two special values of g_l are $g_0 = 0$, which follows from the normalization of $P(\mathbf{s}, \mathbf{s}')$ and $g_1 = c/l_{\mathrm{tr}}$ where l_{tr} is the transport mean free path, defined by $l_{\mathrm{tr}} = c/[\mu_s(1 - \langle\cos\theta\rangle)]$, where $\langle\cos\theta\rangle$ is the average of the cosine of the scattering angle θ. Equation (2.4) serves as the exact Green's function of light propagation in the velocity space. Since in an infinite uniform medium this function is independent of the source position, r_0, the requirements for a Green's function are satisfied. Indeed, in an infinite uniform medium, this propagator determines the particle migration behavior, including its spatial distribution, because displacement is an integration of velocity over time. The distribution function $I(\mathbf{r}, \mathbf{s}, t)$ (the source is located at $\mathbf{r}_0 = 0$) is given by

$$I(\mathbf{r}, \mathbf{s}, t) = \left\langle \delta\left(\mathbf{r} - c\int_0^t \mathbf{s}(t')\mathrm{d}t'\right)\delta(\mathbf{s}(t) - \mathbf{s})\right\rangle \tag{2.6}$$

where $\langle\ldots\rangle$ means the ensemble average in the velocity space. The first delta function imposes that the displacement is given by the path integral. The second delta function ensures the correct final value of direction. Equation (2.6) is an exact formal solution of Eq. (2.1), but cannot be evaluated directly. We make a Fourier transform for the first delta function in Eq. (2.6), and then make a cumulant expansion, defined by Ma (1985)

$$\langle\exp(ikx)\rangle = \exp\left(\sum_{n=1}^{\infty}\langle x^n\rangle_c (ik)^n/n!\right) \tag{2.7}$$

where the first cumulant $\langle x\rangle_{\mathrm{c}} = \langle x\rangle$ is the central position of x, and $\langle x^2\rangle_{\mathrm{c}} = \langle x^2\rangle - \langle x\rangle\langle x\rangle$ is the half-width of the distribution. $\langle x^3\rangle_{\mathrm{c}}$ describes the skewness or asymmetry of the distribution, and $\langle x^4\rangle_{\mathrm{c}}$ describes the kurtosis or the difference from the Gaussian-shaped bell form, and so on. We obtain

$$\begin{aligned} I(\mathbf{r}, \mathbf{s}, t) = F(\mathbf{s}, \mathbf{s_0}, t)\frac{1}{(2\pi)^3}\int \mathrm{d}\mathbf{k}\exp\Bigg\{ i\mathbf{k}\cdot\mathbf{r} + \sum_{n=1}^{\infty}\frac{(-ic)^n}{n!}\sum_{j_n}\cdots\sum_{j_1} k_{j_n}\ldots k_{j_1} \\ \left\langle \int_0^t \mathrm{d}t_n \ldots \int_0^t \mathrm{d}t_1 T\left[s_{j_n}(t_n)\ldots s_{j_1}(t_1)\right]\right\rangle_c \Bigg\} \end{aligned} \tag{2.8}$$

where T denotes time-ordered multiplication. In Eq. (2.8), the subscript "c" denotes the cumulant. As shown above, cumulants are just centered moments. Hence, if moments $\langle \mathrm{A}^m\rangle$ ($m = 1, 2, \ldots, n$) have been calculated, $\langle \mathrm{A}^m\rangle_c$ can be recursively obtained, and vice versa.

In the following, we derive the analytical expression for the moment ensemble average. Using a standard time-dependent Green's function approach, it is given by

$$\left\langle \int_0^t \mathrm{d}t_n \ldots \int_0^t \mathrm{d}t_1 T\left[s_{j_n}(t_n)\ldots s_{j_1}(t_1)\right]\right\rangle = \frac{1}{F(\mathbf{s},\mathbf{s_0},t)}\left\{\int_0^t \mathrm{d}t_n \int_0^{t_n} \mathrm{d}t_{n-1}\ldots \int_0^{t_2} \mathrm{d}t_1 \int \mathrm{d}\mathbf{s}^{(n)} \right.$$
$$\left. \ldots \int \mathrm{d}\mathbf{s}^{(1)} F\left(\mathbf{s},\mathbf{s}^{(n)},t-t_n\right) s_{j_n}^{(n)} F\left(\mathbf{s}^{(n)},\mathbf{s}^{(n-1)},t_n-t_{n-1}\right)\ldots s_{j_1}^{(1)} F\left(\mathbf{s}^{(1)},\mathbf{s}_0,t_1-0\right) + \text{permut.}\right\} \tag{2.9}$$

where the word "permut." means all $n! - 1$ terms obtained by permutation of $\{j_i\}$, $i = 1, \ldots, n$, from the first term. An intuitive way to understand Eq. (2.9) is to use a basic concept in quantum mechanics that the left side of the equation is written in the Heisenberg representation while the right side of the equation is written in the Schrodinger representation. Equation (2.9) can also be understood as a Feynman path integral.

In Eq. (2.9), $F(\mathbf{s}^{(i)}, \mathbf{s}^{(i-1)}, t_i - t_{i-1})$ is given by Eq. (2.4). Since Eq. (2.4) is exact, Eq. (2.9) provides the exact nth order moments of the distribution.

In Cartesian coordinates three components of $\mathbf{s}$ are s_x, s_y, and s_z. For convenience in calculation, however, we will use the components of $\mathbf{s}$ on the base of spherical harmonics:

$$\begin{aligned}\mathbf{s} &= [s_1, s_0, s_{-1}] = [Y_{11}(\mathrm{s}), Y_{10}(\mathrm{s}), Y_{1-1}(\mathrm{s})] \\ &= \left[-\frac{1}{2^{1/2}}\sin\theta \mathrm{e}^{+i\phi}, \cos\theta, \frac{1}{2^{1/2}}\sin\theta \mathrm{e}^{-i\phi}\right]\end{aligned} \tag{2.10}$$

The recurrence relation of the spherical harmonics is given by

$$Y_{lm}(s)Y_{1j}(s) = \sum_i Y_{l+i,m+j}(s)\langle l,1,m,j|l+i,m+j\rangle\langle l,1,0,0|l+i,0\rangle \tag{2.11}$$

where $i = \pm 1$. $\langle l,1,m,j|l+i,m+j\rangle$ is the Clebsch–Gordan coefficients of angular momentum theory (Brink and Satchler 1962). The orthogonality relation of spherical harmonics is given by

$$\int \mathrm{d}s Y^*_{l'm'}(s)Y_{lm}(s) = \frac{4\pi}{2l+1}\delta_{l,l'}\delta_{m,m'} \tag{2.12}$$

Using Eqs. (2.11) and (2.12), the integrals over $\mathrm{d}\mathbf{s}^{(n)} \ldots \mathrm{d}\mathbf{s}^{(1)}$ in Eq. (2.9) can be analytically performed. We obtain, when $\mathbf{s}_0$ is set to along $\mathbf{z}$, that

$$\left\langle \int_0^t \mathrm{d}t_n \ldots \int_0^t \mathrm{d}t_1 T\left[s_{j_n}(t_n)\ldots s_{j_1}(t_1)\right]\right\rangle = \frac{1}{F(s,s_0,t)}\left\{\sum_l Y_{\left[l,\sum_{m=1}^n j_m\right]}(s)\right.$$
$$\sum_{i_n}\cdots\sum_{i_1}\frac{2\left(l-\sum_{m=1}^n i_m\right)+1}{4\pi}\prod_{k=1}^n\left\langle l-\sum_{m=1}^{n-k+1} i_{n-m+1},1,\sum_{m=1}^{k-1} j_m, j_k \middle| l-\sum_{m=1}^{n-k} i_{n-m+1},\sum_{m=1}^{k} j_m\right\rangle$$
$$\left.\left\langle l-\sum_{m=1}^{n-k+1} i_{n-m+1},1,0,0\middle| l-\sum_{m=1}^{n-k} i_{n-m+1},0\right\rangle \quad D^l_{i_n,\ldots,i_1}(t)+\text{permut.}\right\} \tag{2.13}$$

with $i_f = \pm 1$, $f = 1, 2, \ldots, n$, and

$$D^l_{i_n,\ldots,i_1}(t) = \exp(-\mu_a t)\left\{\int_0^t \mathrm{d}t_n \int_0^{t_n} \mathrm{d}t_{n-1}\ldots\int_0^{t_2}\mathrm{d}t_1\right.$$
$$\exp[-g_l(t-t_n)]\exp[-g_{l-i_n}(t_n-t_{n-1})] \tag{2.14}$$
$$\left.\ldots\exp\left[-g_{l-\sum_{k=1}^n i_{n-k+1}}(t_1-0)\right]\right\}$$

Note that all ensemble averages have been performed. Equation (2.14) involves integrals of exponential functions, which can be analytically performed. Equation (2.14) includes all related scattering and absorption parameters, g_l, $l = 0$, 1, …, and μ_a, that determines the time evolution dynamics. The final particle direction, **s**, appears as the argument of the spherical harmonics $Y_{lm}(\mathbf{s})$ in Eq. (2.13). Substituting Eq. (2.14) into Eq. (2.13), and using a standard cumulant procedure, the cumulants as functions of angle **s** and time t up to an arbitrary nth order can be analytically computed. The final position, **r**, appears in Eq. (2.8), and its component can be expressed as $|\mathbf{r}|$ and Y_{1j}, $j = 1, 0, -1$. Then, performing a numerical three-dimensional inverse Fourier transform over **k**, an approximate distribution function, $I(\mathbf{r}, \mathbf{s}, t)$, accurate up to the nth order cumulant, is obtained.

3 Gaussian Approximation of the Distribution Function

By a cutoff at the second-order cumulant, the integral over **k** in Eq. (2.8) can be analytically performed, which directly leads to a Gaussian spatial distribution, which will be displayed in Eq. (3.2). The exact first cumulant provides the correct center position of the distribution. The exact second cumulant provides the correct half-width of spread of the distribution. The expressions below are given in Cartesian coordinates with subscripts α, β = [*x*, *y*, *z*]. These expressions are obtained by use of the following unitary transform $s_\alpha = \sum U_{\alpha j} s_j$, $j = 1, 0, -1$, from Eq. (2.13) (up to the second-order cumulant) which is based on $s_j = Y_{1j}$ (s), with

$$U = \begin{bmatrix} -2^{-1/2} & 0 & 2^{-1/2} \\ 2^{-1/2}i & 0 & 2^{-1/2}i \\ 0 & 1 & 0 \end{bmatrix} \tag{3.1}$$

We set $\mathbf{s_0}$ along the $\mathbf{z}$ direction and denote $\mathbf{s}$ as (θ, ϕ). Our cumulant approximation to the distribution function reduces to

$$I(r,s,t) = \frac{F(s,s_0,t)}{(4\pi)^{3/2}} \frac{1}{(\det B)^{1/2}} \exp\left[-\frac{1}{4}\left(B^{-1}\right)_{\alpha\beta}(r-r^c)_\alpha (r-r^c)_\beta\right], \tag{3.2}$$

with the center of the packet (the first cumulant), denoted by $\mathbf{r}^c$, located at

$$r_z^c = G\sum_l A_l P_l(\cos\theta)[(l+1)f(g_l - g_{l+1}) + lf(g_l - g_{l-1})] \tag{3.3}$$

$$r_x^c = G\sum_l A_l P_l^{(1)}(\cos\theta)\cos\phi[f(g_l - g_{l-1}) - f(g_l - g_{l+1})], \tag{3.4}$$

where $G = c\exp(-\mu_a t)/F(\mathbf{s}, s_0, t)$, $A_l = (1/4\pi)\exp(-g_l t)$, and

$$f(x) = [\exp(xt) - 1]/x \tag{3.5}$$

r_y^c is obtained by replacing cosϕ in Eq. (3.4) by sinϕ. In Eqs. (3.3) and (3.4), $P_l^{(m)}(\cos\theta)$ is the associated Legendre function.

The square of the average spread width (the second-order cumulant) is given by

$$B_{\alpha\beta} = cG\Delta_{\alpha\beta} - r_\alpha^c r_\beta^c/2, \tag{3.6}$$

where all the coefficients are functions of angle and time:

$$\Delta_{zz} = \sum_l A_l P_l(\cos\theta)\left[\frac{l(l-1)}{2l-1}E_l^{(1)} + \frac{(l+1)(l+2)}{2l+3}E_l^{(2)} + \frac{l^2}{2l-1}E_l^{(3)} + \frac{(l+1)^2}{2l+3}E_l^{(4)}\right] \tag{3.7}$$

$$\begin{aligned}\Delta_{xx,yy} = \sum_l \frac{1}{2}A_l P_l(\cos\theta)\Bigg[&-\frac{l(l-1)}{2l-1}E_l^{(1)} - \frac{(l+1)(l+2)}{2l+3}E_l^{(2)} \\ &+ \frac{l(l-1)}{2l-1}E_l^{(3)} + \frac{(l+1)(l+2)}{2l+3}E_l^{(4)}\Bigg] \pm \sum_l \frac{1}{2}A_l P_l^{(2)}(\cos\theta)\cos(2\phi) \\ &\left[\frac{1}{2l-1}E_l^{(1)} + \frac{1}{2l+3}E_l^{(2)} - \frac{1}{2l-1}E_l^{(3)} - \frac{1}{2l+3}E_l^{(4)}\right]\end{aligned} \tag{3.8}$$

where (+) corresponds to Δ_{xx} and (−) corresponds to Δ_{yy},

$$\Delta_{xy} = \sum_l \frac{1}{2} A_l P_l^{(2)}(\cos\theta)\sin(2\phi)\left[\frac{1}{2l-1}E_l^{(1)} + \frac{1}{2l+3}E_l^{(2)} - \frac{1}{2l-1}E_l^{(3)} - \frac{1}{2l+3}E_l^{(4)}\right] \tag{3.9}$$

$$\Delta_{xz} = \sum_l \frac{1}{2} A_l P_l^{(1)}(\cos\theta)\cos(\phi)\left[\frac{2(l-1)}{2l-1}E_l^{(1)} + \frac{2(l+2)}{2l+3}E_l^{(2)} + \frac{1}{2l-1}E_l^{(3)} + \frac{1}{2l+3}E_l^{(4)}\right] \tag{3.10}$$

Δ_{yz} is obtained by replacing cos (ϕ) in Eq. (3.10) by sin (ϕ). In Eqs. (3.7)–(3.10)

$$E_l^{(1)} = [f(g_l - g_{l-2}) - f(g_l - g_{l-1})]/(g_{l-1} - g_{l-2}) \tag{3.11}$$

$$E_l^{(2)} = [f(g_l - g_{l+2}) - f(g_l - g_{l+1})]/(g_{l+1} - g_{l+2}) \tag{3.12}$$

$$E_l^{(3)} = [f(g_l - g_{l-1}) - t)]/(g_l - g_{l-1}) \tag{3.13}$$

$$E_l^{(4)} = [f(g_l - g_{l+1}) - t)]/(g_l - g_{l+1}) \tag{3.14}$$

The photon density $N(\mathbf{r}, t)$ of the second cumulant solution is given by

$$N(\mathbf{r}, t) = \frac{1}{(4\pi D_{zz} ct)^{1/2}} \frac{1}{4\pi D_{xx} ct} \exp\left[-\frac{(z - R_z)^2}{4D_{zz}ct}\right] \exp\left[-\frac{(x^2 + y^2)}{4D_{xx}ct}\right] \exp(-\mu_a t) \tag{3.15}$$

with the mean position

$$R_z = c[1 - \exp(-g_1 t)]/g_1. \tag{3.16}$$

The corresponding time-dependent diffusion coefficients are

$$D_{zz} = \frac{c}{3t}\left\{\frac{t}{g_1} + \frac{3g_1 - g_2}{g_1^2(g_1 - g_2)}[1 - \exp(-g_1 t)] + \frac{2}{g_2(g_1 - g_2)}[1 - \exp(-g_2 t)] - \frac{3}{2g_1^2}[1 - \exp(-g_1 t)]^2\right\} \tag{3.17}$$

$$D_{xx} = D_{yy} = \frac{c}{3t}\left\{\frac{t}{g_1} + \frac{g_2}{g_1^2(g_1 - g_2)}[1 - \exp(-g_1 t)] - \frac{1}{g_2(g_1 - g_2)}[1 - \exp(-g_2 t)]\right\} \tag{3.18}$$

We see that, at very early time, the diffusion coefficient is near zero, and the center of motion moves with speed c. This stage is ballistic-like mode. With increase of time, the motion of center slows down and the diffusion coefficients increase from zero. This stage of photon migration is often called a "snack-like mode." At large times, the center of motion tends toward l_{tr} and diffusion coefficients to $l_{tr}/3$, which is the result of diffusion mode (Figs. 1, 2 , 3, and 4).

The above analytical solution, although it has exact center and half-width, is not satisfactory in two aspects. First, at very early times, $\exp(-g_l t) \rightarrow 1$, for all l; hence, one cannot ensure summation over l to be convergent. Second, particles at front edge of the Gaussian distribution travel faster than speed c, thus violating causality.

In order to make the summation over l convergent, we separate the ballistic component from the total distribution and the compute the cumulants for the scattered component. By use of the Legendre expansion of delta function $\delta(\mathbf{s} - \mathbf{s}_0)$, the first and second cumulant expressions of the scattered component can be easily obtained.

For practical applications, we use a semi-phenomenological model. The Gaussian distribution is replaced by a new shaped distribution, which is given by a Gaussian distribution, multiplied by a factor in order to ensure causality. By fitting the exact center position and half-width of distribution computed by the cumulant solution, the related parameters of the new shaped distribution can be determined. The formulation for the above improvement of cumulant solution is given in Lax et al. (2006), Cai et al. (2005).

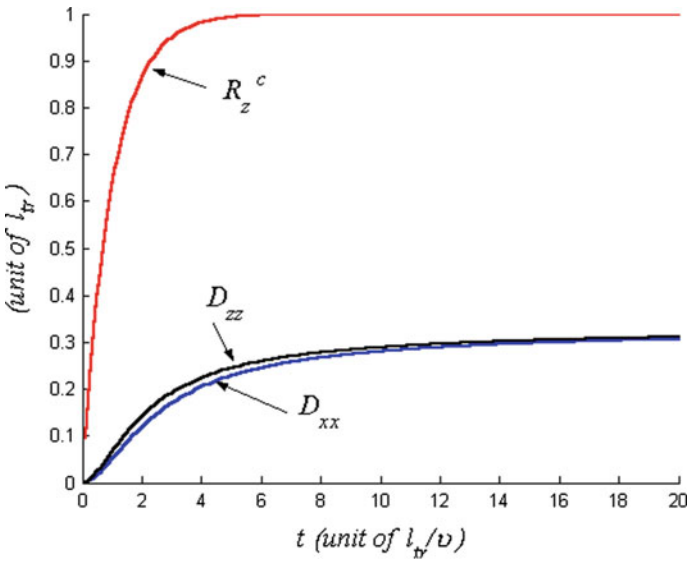


Fig. 1 The moving center of photons, R_z^c, and the diffusion coefficients, D_{zz} and D_{xx}, as function of time, where g_l are calculated by Mie theory, assuming water drops with $a/\lambda = 1$, with a the radius of droplet and λ the wavelength of light, and the index of refraction $m = 1.33$

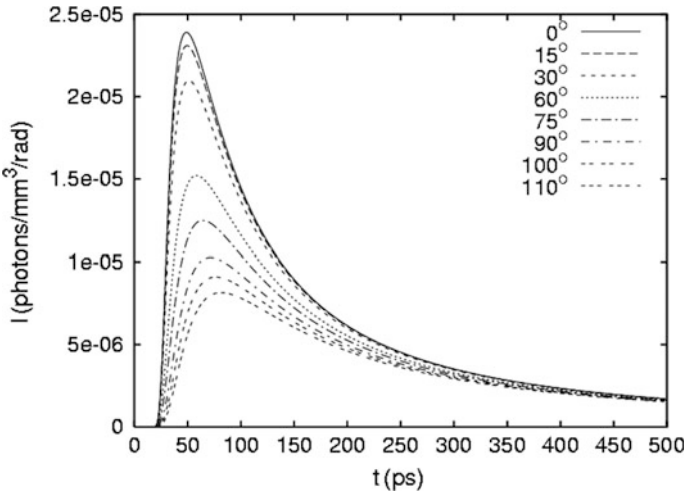


Fig. 2 The time-resolved profile of light at different angles measured on a detector 10 mm from the source in the incident direction. The parameters are $l_{\mathrm{tr}} = 2$ mm, $l_a = 300$ mm, the phase function is computed using Mie theory for polystyrene spheres in water, with diameter $d = 1.11$ μm, and the wavelength of laser source $\lambda = 625$ nm, which gives the g-factor $g = 0.926$

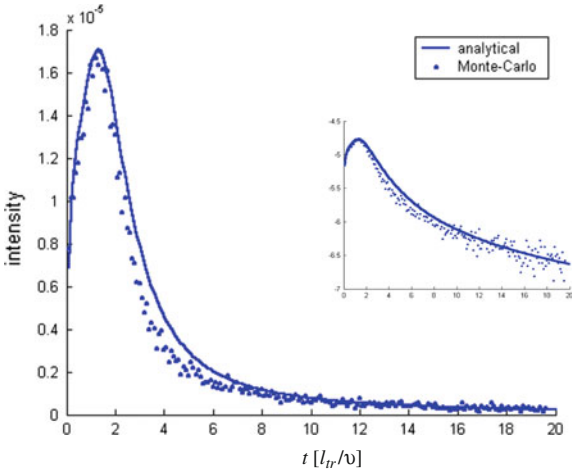


Fig. 3 Time-resolved profile of the backscattered (180°) photon intensity inside a disk with center at $\mathbf{r} = 0$, radius $R = 1l_{\mathrm{tr}}$, thickness $dz = 0.1l_{\mathrm{tr}}$ and the received angle $d\cos\theta = 0.001$, normalized to inject 1 photon. The Heyney–Greenstein phase function with $g = 0.9$ is used, and $1/l_a = 0$. The solid curve is for the second cumulant solution (Gaussian distribution), and dots are for Monte Carlo simulation. The inset diagram is same result drawn using a log scale for intensity

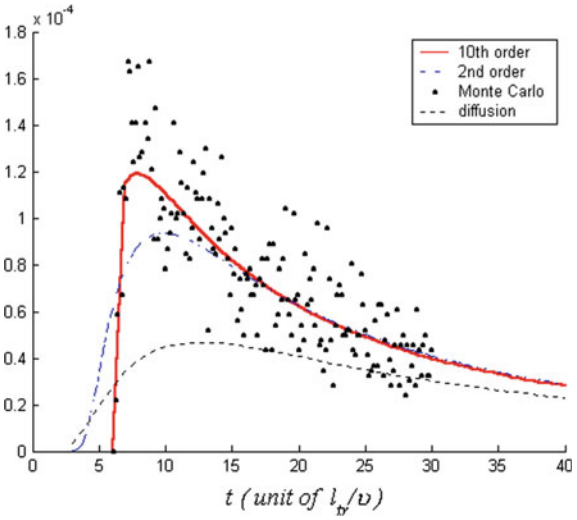


Fig. 4 Time-resolved profile of transmission light in an infinite uniform medium, computed using the tenth order of cumulants solution (*solid curve*), the second cumulant solution (*dotted curve*), and the diffusion approximation (*thick dotted curve*), comparing with that of Monte Carlo simulation (*discrete dots*). The detector is located at $z = 6l_{\mathrm{tr}}$ from source along the incident direction, and received direction is $\theta = 0$. The Heyney–Greenstein phase function with $g = 0.9$ is used, and the absorption coefficient $1/l_a = 0$

4 Cumulant Solution for Polarized Radiative Transport Equation

Understanding polarization phenomenon in a multiple scattering medium is important because of its possible applications in biomedical media, cloud monitoring, and communications. This subject has been investigated for many years since the polarized photon transport equation was derived (Ishimaru 1978; Rybicki 1996).

In Stokes representation (SP), the polarized light is described by vector $\mathbf{I}^{\mathrm{SP}} = [I, Q, U, V]$. We will use an circular representation (CP): the parallel-polarized component $I_{\parallel} = (I + Q)/2$, the perpendicular-polarized component $\mathrm{I}\perp = (I - Q)/2$, the right-handed circular polarized component $I_R = (I + V)/2$, and the left-handed circular polarized component $I_L = (I - V)/2$.

The transport equation for the polarized photon distribution function $\mathbf{I}(\mathbf{r},\mathbf{s},t)$ in an infinite uniform medium, from a point pulse light source, $\mathbf{I}^{(0)}\delta(\mathbf{r} - \mathbf{r}_0)\delta(\mathbf{s} - \mathbf{s}_0)\delta(t - 0)$ is given by

$$\begin{aligned} &\partial I^{\mathrm{SP}}(r,s,t)/\partial t + cs \cdot \nabla_r I^{\mathrm{SP}}(r,s,t) + \mu_a I^{\mathrm{SP}}(r,s,t) = \mu_s \int P(s,s') \\ &\left[I^{\mathrm{SP}}(r,s',t) - I^{\mathrm{SP}}(r,s,t)\right]\mathrm{d}s' + I^{\mathrm{SP}^{(0)}}\delta(r - r_0)\delta(s - s_0)\delta(t - 0) \end{aligned} \tag{4.1}$$

where vector $\mathbf{I}(\mathbf{r}, \mathbf{s}, t)$ has four polarization components, c is the light speed in the medium, μ_s is the scattering rate, μ_a is the absorption rate, and $\mathbf{P}(\mathbf{s}, \mathbf{s}')$ is the 4×4 phase matrix. A meridian plane parallel to the z-axis and the light direction $\mathbf{s}$ is used as plane of reference for the description of the polarization state. With a rotation of

reference plane through an angle $\alpha \geq 0$ (in the counterclockwise direction, when looking in the direction of propagation) $\mathbf{I}^{\mathrm{SP}}$ varies as

$$\left(\mathbf{I}^{\mathrm{SP}}\right)' = L^{\mathrm{SP}}(\alpha)\mathbf{I}^{\mathrm{SP}} \tag{4.2a}$$

This relation is given by

$$\begin{bmatrix} I' \\ Q' \\ U' \\ V' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\alpha & \sin 2\alpha & 0 \\ 0 & -\sin 2\alpha & \cos 2\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} \tag{4.2b}$$

One has to be careful that all the Stokes parameters are referred to the same reference plane of the propagation direction. Therefore, a double multiplication by the transformation matrix L is necessary, first before scattering to refer the incident light to the scattering plane, then after scattering to change from the scattering plane to the meridian plane of the scattering direction. With fixed coordinates, zenith angle θ and azimuth ϕ, the phase matrix for scattering from a direction (θ', ϕ') to a direction (θ, ϕ) is given by

$$\mathbf{P}(\mathbf{s}, \mathbf{s}') = \mathbf{L}(\pi - \chi)\mathbf{P}(\cos\Theta)\mathbf{L}(-\chi') \tag{4.3}$$

where χ' and χ are the two rotation angles. Θ is the scattering angle between light rays before $[\mathbf{s}(\theta,\phi)]$ and after scattering $[\mathbf{s}'(\theta', \phi')]$. The matrices $\mathbf{L}(-\chi')$ and $\mathbf{L}(\pi - \chi)$ are those required to rotate meridian planes before and after scattering onto or from a local scattering plane, as shown in Fig. 5. The intrinsic property of scattering is described by the 4×4 scattering matrix $\mathbf{P}(\cos\Theta)$, which is assumed to be only dependent on $\cos\Theta = \mathbf{s} \cdot \mathbf{s}'$.

In the CP, $\mathbf{I}^{\mathrm{CP}} = [I_2, I_0, I_{-0}, I_{-2}]$. The relation between $\mathbf{I}^{\mathrm{CP}}$ and $\mathbf{I}^{\mathrm{SP}}$ is given by

$$\mathbf{I}^{\mathrm{CP}} = T\mathbf{I}^{\mathrm{SP}} \tag{4.4a}$$

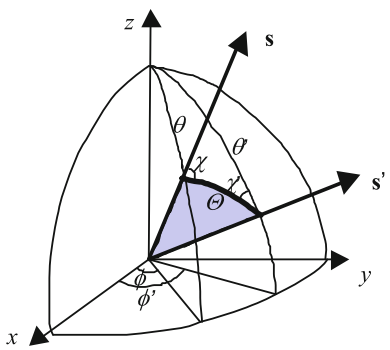


Fig. 5 Geometry of the scattering plane and the reference planes related to the incident ray, $\mathbf{s}'(\theta', \phi')$, and the scattered ray, $\mathbf{s}(\theta, \phi)$. The dark plane is the scattering plane. χ is the angle between the meridian plan $(\mathbf{s}, \mathbf{z})$ and the scattering plane. χ' is the angle between the meridian plane $(\mathbf{s}', \mathbf{z})$ and the scattering plan

$$T = \frac{1}{2}\begin{bmatrix} 0 & 1 & i & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -i & 0 \end{bmatrix} \quad \text{and} \quad T^{-1} = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ -i & 0 & 0 & i \\ 0 & 1 & -1 & 0 \end{bmatrix} \tag{4.4b}$$

In CP a rotation of the reference plane through an angle α around the light direction can now be written as

$$I'^{\mathrm{CP}} = L^{\mathrm{CP}}(\alpha)I^{\mathrm{CP}} = TI'^{\mathrm{SP}} = TL^{\mathrm{SP}}(\alpha)I^{\mathrm{SP}} = TL^{\mathrm{SP}}(\alpha)T^{-1}I^{\mathrm{CP}} \tag{4.5a}$$

where

$$L^{\mathrm{CP}}(\alpha) = TL^{\mathrm{SP}}(\alpha)T^{-1} = \begin{bmatrix} \mathrm{e}^{-i2\alpha} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \mathrm{e}^{i2\alpha} \end{bmatrix} \tag{4.5b}$$

For the phase matrix, transform between two representations is $\mathbf{P}^{\mathrm{CP}} = \mathbf{T}\mathbf{P}^{\mathrm{SP}}\mathbf{T}^{-1}$.

The elements of the scattering matrix for polarized light can be expanded in generalized spherical functions (GSF) (Brink and Satchler 1962):

$$P_{mn}(\mu, \phi; \mu', \phi') = \frac{1}{4\pi}\sum_{l} p^{l}_{mn} \sum_{s=-l}^{l} (-1)^{s} P^{l}_{m,s}(\mu) P^{l}_{s,n}(\mu') \exp[-is(\phi - \phi')] \tag{4.6}$$

with m, n = 2, 0, −0, −2, and $l \geq \max(|m|, |n|)$. The functions $P^{l}_{m,n}(\cos\Theta)$ are connected to Wigner d-function in angular momentum as

$$d^{j}_{m,n}(\cos\Theta) = i^{m-n} P^{j}_{m,n}(\cos\Theta), \quad 0 \leq \Theta < \pi \tag{4.7}$$

The coefficients p^{l}_{mn} in Eq. (4.6) provide an intrinsic description of the scattering mechanism, with properties $p^{l}_{mm} = p^{l}_{m-m}$ are real; $p^{l}_{mn} = p^{l}_{nm} = p^{l}_{-m-n}$; $p^{l}_{20} = \left[p^{l}_{2-0}\right]^{*}$. Therefore, for each $l \geq 2$, there are six independent real elements: $p^{l}_{00}, p^{l}_{22}, p^{l}_{0-0}, p^{l}_{2-2}, \mathrm{Re}[p^{l}_{20}]$, and $\mathrm{Im}[p^{l}_{20}]$. For l = 0, 1, only p^{l}_{00} and p^{l}_{0-0} are nonzero.

Similar to the cumulant solution in Sect. 3, the analytical solution of time-dependent polarized photon transport equation (4.1) can be solved (Cai et al. 2000b, 2006). Dynamics of the photon distribution in the light direction space on a spherical surface of radius 1, $\mathbf{F}(\mathbf{s}, \mathbf{s}_0, t)$, is first studied. The mth component of the $\mathbf{F}(\mathbf{s}, \mathbf{s}_0, t)$ in CP, with the initial polarization in unit n_0 state, has the following exact form:

$$F_{mn_0}(\mathbf{s},\mathbf{s}_0,t)=\frac{1}{4\pi}\sum_l F^l_{mn_0}(t)\sum_{s=-l}^{l}(-1)^s P^l_{m,s}(\mu)P^l_{s,n}(\mu_0)\exp[-is(\phi-\phi_0)\exp[-\mu_a t]. \tag{4.8}$$

In Eq. (4.8), $F^l_{mn_0}(t)$ obeys an analytically solvable equation:

$$\mathrm{d}F^l_{mn_0}(t)/\mathrm{d}t=-\sum_n \Pi^l_{mn}F^l_{nn_0}(t),\quad \text{with } \Pi^l_{mn}=\mu_s\left[\delta_{m,n}-p^l_{mn}/(2l+1)\right].$$

The solution has the following form: $F^l_{mn_0}(t)=\sum_i [B^l_{mn_0}]_i \exp(-\lambda_i t)$, where the eigenvalues, λ_i, and the constant coefficients, $[B^l_{mn_0}]_i$, can be analytically determined using standard linear algebra from the initial condition: $F^l_{mn_0}(0)=\delta_{m,n_0}(2l+1)/4\pi$. Equation (4.8) provides an exact CP propagator in the light direction space.

By a cutoff at the second cumulant order, a Gaussian analytical approximation expression of the polarized photon spatial distribution is given by

$$I_v(r,s,t)=\frac{F_v(s,t)}{(4\pi)^{3/2}}\frac{1}{[\det D_v]^{1/2}}\exp\left\{-\frac{1}{4}\left[(D_v)^{-1}\right]_{\alpha\beta}\left(r_\alpha-R^c_{v,\alpha}\right)\left(r_\beta-R^c_{v,\beta}\right)\right\} \tag{4.9}$$

where $v=||,\perp,R,L$, representing parallel, perpendicular, right-handed circularly (RH), and left-handed circularly (LH) polarized light, respectively. The factor of the exponential function is summation with α, β = x, y, z. In Eq. (4.9), $r_\alpha = x, y, z$ is the coordinate of the detector. **s** is the direction of the detected scattered light. F_v(**s**, t) in Eq. (4.8) is the angular distribution of scattered light, and $R^c_{v,\alpha}$ represents the position of the average center of the distribution, which is given by $\langle R\rangle$ in Eq. (4.10); $D_v = \langle RR\rangle - \langle R\rangle\langle R\rangle$ is related to the half-width of the spread of the distribution, with $\langle RR\rangle$ given by Eq. (4.11). The subindex is, separately, related to the parallel ($v=||$), perpendicular ($v=\perp$), right-handed circularly ($v=R$), and left-handed circularly ($v=L$) polarized scattered light. Using a standard Green's function approach and a recurrence relation of GSF, $\langle \mathbf{R}_j\rangle$ and $\langle \mathbf{R}_{j_1}\mathbf{R}_{j_2}\rangle$ in CP are obtained as

$$\langle R_j\rangle_{mn_0}=c\sum_l \frac{4\pi}{2l+1}P^l_{m,n_0-j}(\cos\theta)\mathrm{e}^{-i(n_0-j)\phi}\gamma_j\sum_n\sum_k D^{l,k}_{m,n,n_0}(t)C^{l,n}_{l-k,1,n,0}C^{l,n_0-j}_{l-k,1,n_0,-j} \tag{4.10}$$

where n = 2, 0, −0, −2, k = +1, 0, −1; γ_0 = 1, $\gamma_{+1} = -2^{1/2}i$, $\gamma_{-1} = 2^{1/2}i$, $C^{l,m}_{l_1,l_2,m_1,m_2}$ is the Clebsch–Gordan coefficient in the angular momentum theory, and $D^{l,k}_{m,n,n_0}(t)=\int_0^t \mathrm{d}t' F^l_{mn}(t-t')F^{l-k}_{nn_0}(t')\exp(-\mu_a t)$, which can be integrated analytically.

$$\left\langle R_{j_1} R_{j_2} \right\rangle_{mn_0} = c^2 \sum_l \frac{(4\pi)^2}{2l+1} P^l_{m,n_0-j_1-j_2}(\cos\theta) \mathrm{e}^{-i(n_0-j_1-j_2)\phi} \gamma_{j_1} \gamma_{j_2}$$
$$\sum_{n_1} \sum_{n_2} \sum_{k_1} \sum_{k_2} E^{l,k_1,k_2}_{m,n_1,n_2,n_0}(t) \frac{1}{2(l+k_1)+1} C^{l,n_1}_{l-k_1,1,n_1,0} \tag{4.11}$$
$$C^{l,n_0-j_1-j_2}_{l-k_1,1,n_0-j_2,-j_1} C^{l-k_1,n_2}_{l-k_1-k_2,1,n_2,0} C^{l-k_1,n_0-j_2}_{l-k_1-k_2,1,n_0,-j_2},$$

with $E^{l,k_1,k_2}_{m,n_1,n_2,n_0}(t) = \int_0^t \mathrm{d}t' \int_0^{t'} \mathrm{d}t'' F^l_{mn_1}(t-t') F^{l-k_1}_{n_1 n_2}(t'-t'') F^{l-k_1-k_2}_{n_2 n_0}(t'') \exp(-\mu_a t)$.

As an example, the depolarization of circularly polarized light in a turbid medium is studied (Cai et al. 2006). The incident wavelength in calculations is set as 610 nm. The sample is prepared with polystyrene particles (refraction index is 1.59) suspended in water. The results in the following paragraphs have been scaled to use the transport mean free path l_{tr} as the unit of length and the flight time for one transport mean free path in the medium l_{tr}/c as the unit of time (for example, if l_{tr} = 1 cm, l_{tr}/c corresponds to 33.3 ps).

Figure 6a–d shows the particle size evolution of the temporal profile of the backscattered right-handed (solid curve) and left-handed (dotted curve) circularly polarized light calculated using the polarized cumulant solution. For small size of particles (0.1 μm), the backscattered light is dominated with left-handed circularly polarized light, and the helicity is flipped in the backscattered light (Fig. 6a). With the increase of particle size (0.213 μm), the intensity difference between right- and left-handed circularly backscattered light reduced (Fig. 6b). For large size of particle (0.855, 8.0 μm), however, the helicity is maintained and the backscattered light is dominated with right-handed circularly polarized light.

Figure 7 shows the particle size dependence of right-handed (solid line) and left-handed (dotted line) backscattered (180°) photon intensity for a right-handed circularly polarized incident light. The intensity is obtained from the peak value of time-resolved profiles.

The analytical cumulant solution of polarized light shows the helicity flipped in the scattering of circularly polarized light from a suspension of small particles and helicity preserved from a suspension of large particles. The depolarization of a well-defined incident polarization state from a medium depends on the sizes of the scatterers. For small particles in a medium (particle diameter $a <$ wavelength λ), the transport mean free path (l_{tr}) approximately equals to the scattering mean free path (l_{s}) and the scattering is isotropic. Photons in the backscattering light are those suffering one or few scattering events, thus their circular electric field remains when photons bounce back, leading to the flip of helicity. While for large particles ($a \geq \lambda$), $l_{\mathrm{tr}} > l_{\mathrm{s}}$ the scattering is highly forward and anisotropic; the depolarization length for a circularly polarized light is larger than corresponding length for an incident linearly polarized light (Xu and Alfano 2005a, b). With the increase of particle size, backscattering results from the accumulated effect of many

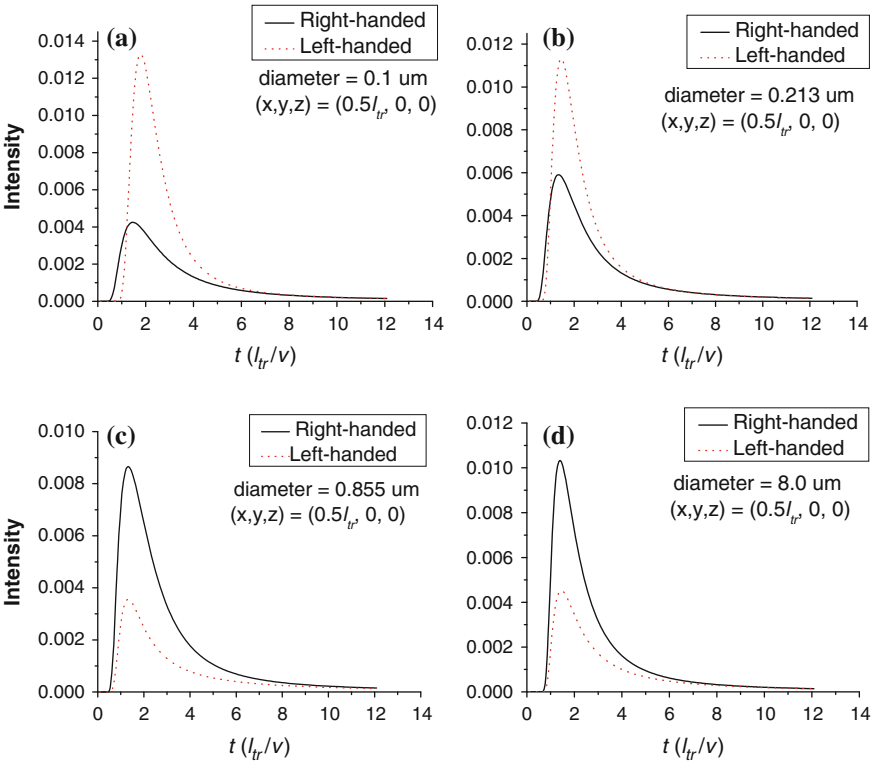


Fig. 6 Particle size dependence of time-resolved profile of the right-handed (*solid*) and left-handed (*dotted*) backscattered (180°) photon intensities at position (x, y, z) = (0.5l_{tr}, 0, 0) inside an infinite medium containing particles of different diameters: 0.1 μm (**a**), 0.213 μm (**b**), 0.855 μm (**c**), 8.0 μm (**d**). The source is incident along the positive z-axis at the origin of coordinate (x, y, z) = (0, 0, 0) and at time zero. The initial Stokes parameter I = [1 0 0 1] represents input right-handed circularly polarized light. The time is normalized to a unit l_{tr}/c

small-angle near-forward scattering events, each of which changes the direction only slightly. These small-angle scattering events do not change the helicity of circularly polarized light. Hence, the backscattered light is dominated by the component that maintains the original helicity. This is known as the "memory effect" in scattering of polarized light (Xu and Alfano 2005a).

In Fig. 8 we compared experimental results of backscattering from large particle suspensions performed by our experimental group with theoretical calculation based on the cumulant solution (Cai et al. 2006).

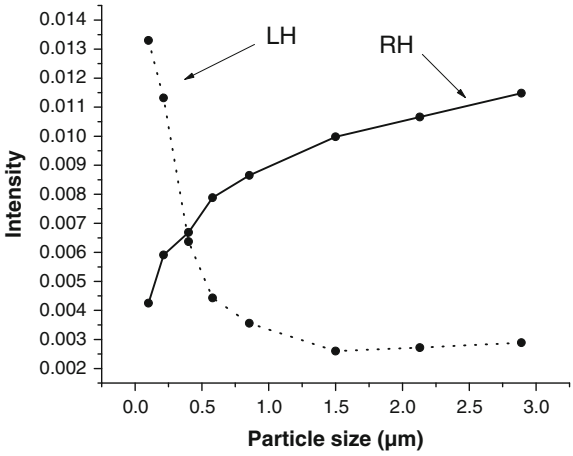


Fig. 7 Particle size dependence of right-handed (*solid line*) and left-handed (*dotted line*) backscattered (180°) photon intensity at position $(x, y, z) = (0.5l_{tr}, 0, 0)$ in an infinite medium. The source is incident along the positive z-axis at the origin of coordinate $(x, y, z) = (0, 0, 0)$. The initial Stokes parameter I = [1 0 0 1] represents input right-handed circularly polarized light

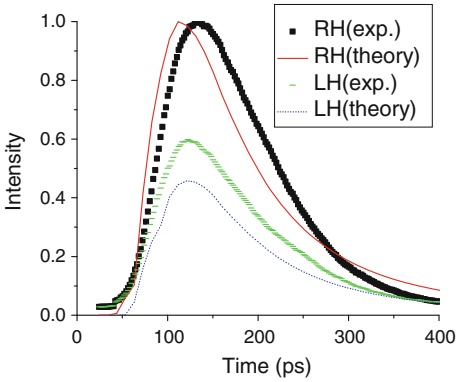


Fig. 8 Comparison of theoretical and experimentally measured normalized intensity distribution of backscattered light. The incident light is right-handed circularly polarized light. The *square* and *dash symbols* refer to experimentally measured right-handed (*RH*) and left-handed (*LH*) circularly polarized light, respectively. The *solid* and *dot curves* are calculated right-handed (*RH*) and left-handed (*LH*) circularly polarized light. The scattering parameters are diameter $d = 8$ μm and scattering coefficient $\mu_s = 1.54\ \text{cm}^{-1}$

5 Applications of the Cumulant Solution of Radiative Transfer

5.1 Early Photon Tomography (EPT)

Niedre et al. (2008) have developed early photon technology for performing high-fidelity fluorescence tomography in living tissues. The theoretical model of the analytical cumulant solution (Eqs. 3.15–3.18) is applied in this subject, which allowed accurate modeling of light propagation in tissue at early time.

EPT uses early arriving photons, that is, photons emitted from an ultrafast laser source (i.e., pulse width <1 ps) which propagate through tissue and are the first to arrive at time-gated detector placed at a distance away from source. The scattering of these early photons is strongly biased in the forward direction and correspondingly they experience a lower number of total scattering events. As a result, they preferentially propagate along significantly less-diffusive paths connecting the source and the detector versus ungated photons and can therefore be used to significantly improve image resolution (Figs. 9 and 10).

5.2 Retrieving Parameters of Water Cloud from CALIPSO Data

The balance of the earth's radiation budget depends on various factors, including microphysical properties of cloud, such as density and size distribution of droplets in cloud, and the thickness of the cloud layer. The active measurement, pursued by the NASA CALIPSO mission, uses laser light pulses from lidars aboard a satellite to probe the atmosphere, and collects the time-resolved *backscattered* signal from each single measurement. Since being launched in 2006, CALIPSO has generated huge amounts of data, which are temporal profiles of 532- and 1064-nm pulses of light backscattered from earth atmosphere. The size and density of scattering entities in earth atmosphere determine the shape and intensity of the temporal profile of backscattered pulses, especially, for early arriving photons.

Difficulty in developing approaches to retrieve dense water cloud parameters from optical data stems from the paucity of adequate theoretical formalism to account for multiple scattering of photons that is prevalent given the high concentration of randomly distributed water droplets inside dense cloud. The approaches (Kokhanovsky 2005; Baum et al. 2000; Nakajima and King 1990; Kokhanovsky et al. 2013) to address the problem of multiple scattering may include numerical solutions of the radiative transfer equation or Monte Carlo simulation of photon migration in cloud. However, these are cumbersome, computation time intensive, and are not suitable as forward models for retrieving the cloud parameters from measured data.

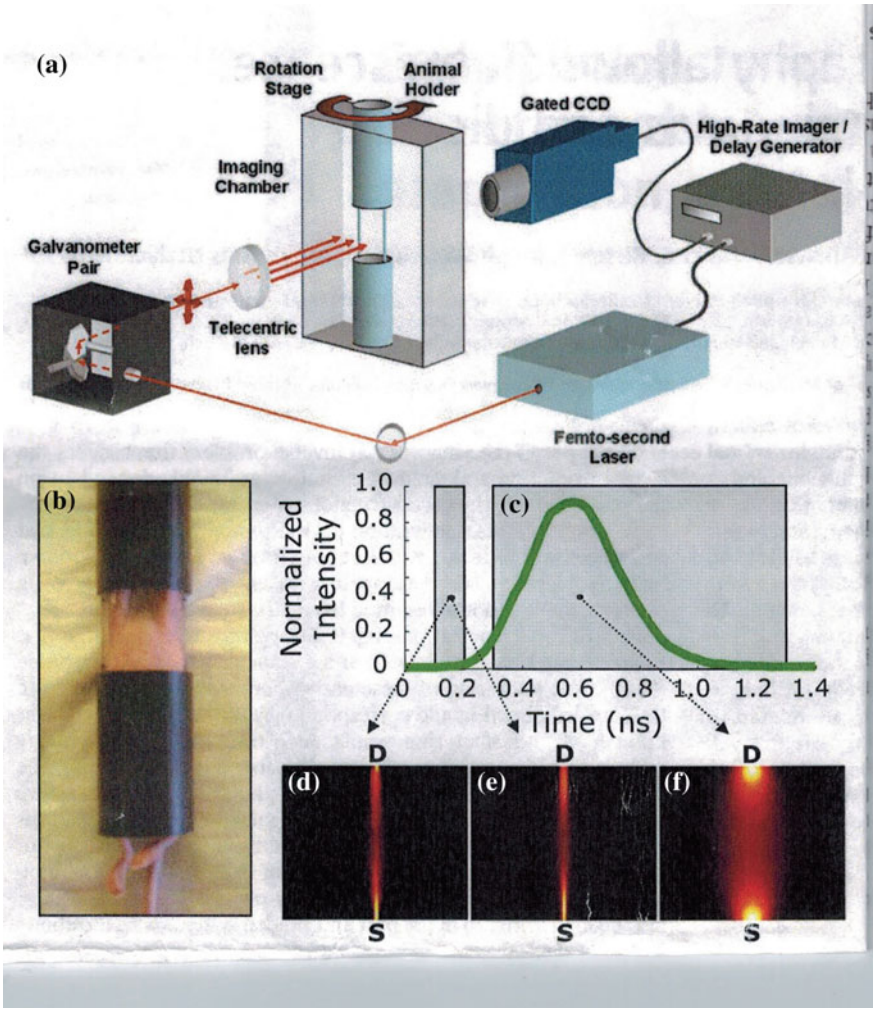


Fig. 9 The early photon tomography (EPT) approach. **a** Experimental setup. Early arriving photons are detected using the gated-intensified CCD system. The animal is placed in the custom built cylindrical carbon fiber tube (**b**) for scanning. As photons propagate through the diffusive medium, they disperse temporally (**c**). The experimentally measured Green's function describing the path of photons propagating between a source (*S*) and detector (*D*) pair, measured at early time gate (**d**) and the early photon forward model used in the reconstructions (**e**). This model was calculated using a normalized cumulant approximation to the Boltzmann transport equation and agreed well with measured weight function. The weight function for diffuse (CW) photons for the same geometry is also shown for comparison (**f**). The radial spread of early arriving photons is drastically reduced, thereby allowing the excellent resolution possible with the EPT [Adapted from Fig. 1 of Niedre et al. (2008)]

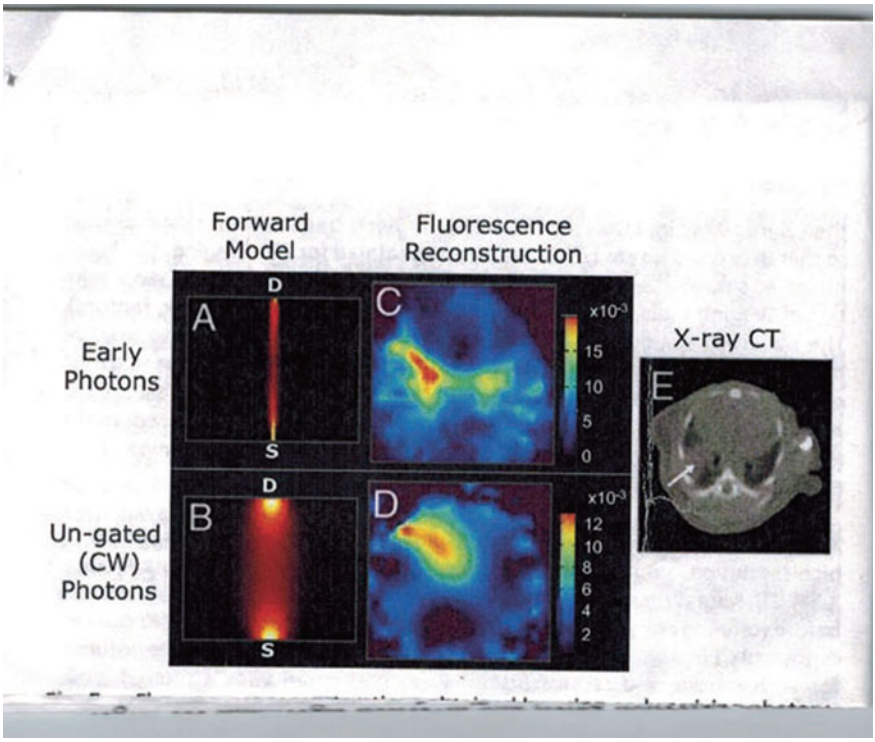


Fig. 10 Fluorescence reconstructions obtained using early arriving photons compared with ungated (continuous wave) photons in mouse with an LLC tumor, 10 days after injection. Forward model for early arriving photons and ungated photons are shown (**a** and **b**), as well as the fluorescence reconstructions for the axial slice (**c** and **d**), and the corresponding X-ray CT axial slice (**e**). The use of early photons enables fluorescence reconstructions with superior resolution and localization of the activated fluorophore in the right lobe of the lung, as well as on the contralateral side compared with the ungated reconstruction [Adapted from Fig. 5 in Niedre et al. (2008)]

The analytical solution of radiative transfer equation based on the cumulant expansion provides a tool, which can compute the photon density distribution function rapidly and accurately taking full account of multiple scattering events. For the backscattering case, our results are shown to be in good agreement with Monte Carlo simulation results.

The phase function of single scattering of light from a water drop, determined by Mie formula, is related to a/λ, where a is the radius of the drop and λ is the wave length of light. The cumulant solution can be used as a forward model for retrieving the density and the average size of water drops in cloud from CALIPSO data (Cai and Gayen 2010).

In a finite medium with boundaries, photons that pass through the boundary may not go back into medium again, while in an infinite medium photons can be scattered back into the region. Only those photons that leak out scatter back into the finite medium from outside, and are scattered out again, account for the difference between the solution in an infinite medium and that in a finite medium. The boundary condition, hence, mainly affects the photon distribution at late times, not at early time period.

An approximate method, similar to extending the solution of the diffusion equation from an infinite medium to semi-infinite and slab geometries, is used (Lax et al. 1987). The technique involves adding image sources to satisfy the "extrapolated" boundary condition. For semi-infinite geometry, it requires the distribution of photons to be zero at an extrapolated plane, $z = -z_c$ with $z_c = \alpha l_{\mathrm{tr}}$, and $\alpha \sim 0.7$ for the vacuum boundary. The physical interface is at $z = 0$. A virtual negative source is added to the original source at $z = -2\alpha l_{\mathrm{tr}}$. At early times, the contribution from the virtual source is negligible. The boundary effect becomes essential at later times. Comparison with Monte Carlo simulation shows that this approach of extension to the semi-infinite geometry is quantitatively sound (Xu et al. 2001).

In the CALIPSO mission, light backscattering data is collected in the following manner. The satellite flies in an orbit at a height of $\sim$700 km above the earth. The lidars on CALIPSO use linearly polarized laser beams to probe the atmosphere. Nanosecond-duration laser light pulses of wavelength 532 and 1064 nm are injected when satellite flies through different positions of earth. The field of view (FOV) of 130 μrad corresponds to an area of radius $R_b \approx 0.05$ km on the surface of cloud, from which backscattering signal is collected, and will be referred to as "footprint" area in subsequent discussion. For each laser pulse, a group of L1B data, which is the total calibrated attenuated backscattering rate β_{i} (in unit of $\mathrm{km}^{-1}\ \mathrm{SR}^{-1}$) versus altitude h_{i} (in unit of km), is collected. Data represent the backscattered intensity of light collected through the footprint area of cloud surface as a function of time (in unit of μsec). Time scale is obtained by mapping the change of altitudes to the round-trip time $\Delta t = 2\Delta h/c$, with c the speed of light.

Figure 11 shows image of CALIPSO data when satellite passed through North America on August 7, 2007. Each of 56,295 horizontal positions represents a location on the earth by its longitudinal and latitudinal coordinates, while each of the 583 vertical positions represents the altitude measured from the earth surface. Strength of the attenuated backscattering (in unit of $\mathrm{km}^{-1}\ \mathrm{SR}^{-1}$) is represented by pseudo color according to the scaled color bar shown in the figure. Each of these vertical positions corresponds to a point on the temporal profile data based on the $\Delta t = 2\Delta h/c$ mapping. The white-colored points represent the strongest signals backscattered from dense water cloud, while the white-gray-colored points represent the weaker signals from rarer cloud. These signals are several orders of magnitude stronger than the ambient background and signals from other particles in

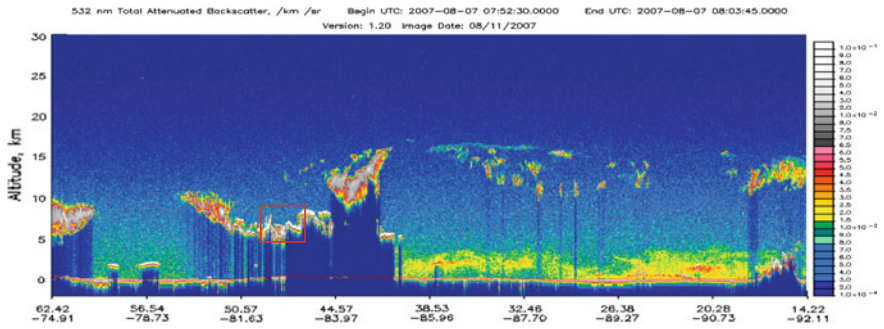


Fig. 11 Image of CALIPSO L1B data is shown when satellite passed through North America on August 7, 2007. The *horizontal points* represent positions on the earth by longitudinal and latitudinal coordinates. The *vertical levels* represent altitude counted from the earth surface. Strength of the attenuated backscattering rate (in unit of km^{-1} SR^{-1}) is represented by pseudo color on the image map, according to the scale shown in the *color bar*

the earth atmosphere. The dark blue color under cloud indicates the strong decay of the tail part of time-resolved signals. The numerical values of measured temporal profiles may be derived from the corresponding hdf data file.

A piece of cloud shown inside the square frame in Fig. 11 is chosen for fitting data to obtain the parameters of cloud. Figure 12 shows some examples of fitting. In Fig. 12a–c the points represent CALIPSO data for the wavelength $\lambda = 532$ nm, and the solid curves are theoretical fits using the analytical model. Table 1 lists the parameters corresponding to Fig. 12a–c at different positions on earth, including satellite's flying UTC time, the longitudinal and latitudinal positions, and the fitted parameters of cloud: the transport mean free path l_{tr}, the effective radius of water drops r_d, the anisotropy factor g, the scattering cross section σ_s, the scattering length l_s, the density of drops ρ, and the altitude of top level of major cloud h_{top}, respectively. We would like to emphasize that the fitting here is against the time-resolved light reflectance, whereas most existing cloud remote sensing algorithms use the continuous wave (CW) measurements. The retrieved effective particle size in the given examples appears to be smaller than the values reported in the literature based on CW measurements (Lin et al. 1998; Wang et al. 2009). This difference may possibly originate from the fact that the particle size recovered from the time-resolved reflectance represents most closely the value at the top of the cloud and the other regions of the cloud have little effects on its value. Remote sensing with time-resolved measurements in principle can yield more accurate characterization of the cloud. Further investigation is needed to resolve the above discrepancy.

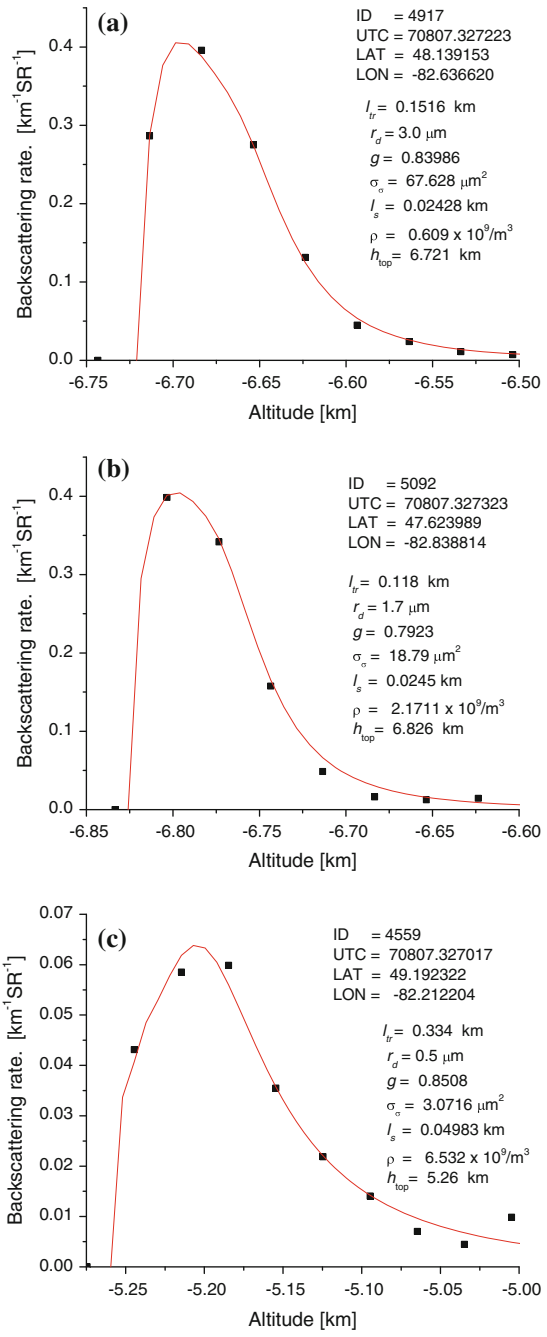


Fig. 12 **a–c** Fitting of CALIPSO data recorded using 532-nm light (represented by *black squares*) to the theoretical formalism (shown by *solid line*) at three different positions on earth for extraction of cloud parameters. The parameters are listed in Table 1, and indicated on respective figures

Table 1 Parametre list using a fitting of CALIPSO data recorded using 532-nm light

Fig.	12(a)	12(b)	12(c)
UTC time	70807.327223	70807.327323	70807.327017
Latitude (°)	48.139153	47.623989	49.192322
Longitude (°)	−82.636620	−82.838814	−82.212204
l_{tr} (km)	0.1516	0.118	0.334
r_d (μm)	3.0	1.7	0.5
g-factor	0.8397	0.7923	0.8508
σ_s (μm^2)	67.628	18.79	3.0716
l_s (km)	0.02428	0.0245	0.04983
ρ ($10^9/m^3$)	0.609	2.171	6.532
h_{top} (km)	6.721	6.826	5.260

6 Summary

In this chapter, we have reviewed the analytical solution of the time-dependent scalar and vector RTE in an infinite uniform medium with an arbitrary light scattering phase function using cumulant expansion. Analytical expressions have been derived for the exact distribution in angle and the spatial cumulants at any angle, exact up to an arbitrary high order. Simple Gaussian analytical approximate expressions of the scalar and vector photon spatial distribution have also been obtained, whose center position and half-width are always exact at arbitrary time. This provides a clear analytical picture that the center of photon distribution advances and the half-width grows in time, showing the evolution of the particle migration from near ballistic, through snake-like, and into the final diffusive regime. The cumulant solution has one unique advantage comparing with other approximation techniques of solving RTE that truncation of the cumulant expansion at order n is exact at that order and cumulants up to and including order n remain unchanged when contributions from higher orders are added. The cumulant solution to RTE is particularly suited to understand and model the temporal migration of scalar and vector photons inside a turbid medium. Two applications of the cumulant solution in Early Photon Tomography (EPT) and cloud sensing from lidar data have been discussed as examples. We expect the cumulant solution will find more applications in biophotonics for optical imaging and remote sensing with time-resolved data in general.

References

Aronson R (1995) Boundary conditions for diffusion of light. J Opt Soc Am A Opt Image Sci Vis 12(11):2532–2539

Baum BA et al (2000) Remote sensing of cloud properties using MODIS airborne simulator imagery during SUCCESS: 1. Data and models. J Geophys Res 105(D9):11767

Brink DM, Satchler GR (1962) Angular momentum. Clarendon Press, Oxford

Cai W, Gayen SK (2010) Retrieval of water cloud characteristic from active sensor data using the analytical solution of radiative transfer equation. J Quant Spectrosc Radiat Transfer 111 (1):226–233

Cai W, Lax M, Alfano R (2000a) Cumulant solution of the elastic boltzmann transport equation in an infinite uniform medium. Phys Rev E 61(4 Pt A):3871–3876

Cai W, Lax M, Alfano RR (2000b) Analytical solution of the polarized photon transport equation in an infinite uniform medium using cumulant expansion. Phys Rev E 63(1):016606

Cai W et al (2002) Diffusion coefficient depends on time, not on absorption. Opt Lett 27(9):731–733

Cai W, Xu M, Alfano RR (2003) Three-dimensional radiative transfer tomography for turbid media. IEEE J Sel Top Quantum Electron 9(2):189–198

Cai W, Xu M, Alfano RR (2005) Analytical form of the particle distribution based on the cumulant solution of the elastic Boltzmann transport equation. Phys Rev E 71(4):041202

Cai W et al (2006) Analytical cumulant solution of the vector radiative transfer equation investigates backscattering of circularly polarized light from turbid media. Phys Rev E 74 (5):056605

Chandrasekhar S (1960) Radiative transfer. Dover, New York

Cuccia DJ et al (2009) Quantitation and mapping of tissue optical properties using modulated imaging. J Biomed Opt 14(2):024012

Ishimaru A (1978) Wave propagation and scattering in random media, vol I a. Academic, New York

Kokhanovsky AA (2005) Satellite-based retrieval of ice cloud properties using a semianalytical algorithm. J Geophys Res 110(D19):D19206

Kokhanovsky AA et al (2013) The determination of cloud optical thickness and effective particle size from measurements of transmitted solar diffuse light. IEEE Geosci Remote Sens Lett 10 (6):1512–1516

Lax M, Nayaramamurti V, Fulton RC (1987) Classical diffusion photon transport in a slab. In: Birman JL, Cummins HZ, Kaplyanskii AA (eds) Laser optics of condensed matter. Plenum, New York, pp 229–237

Lax M, Cai W, Xu M (2006) Random processes in physics and finance. Oxford University Press, Oxford

Liemert A, Kienle A (2012) Infinite space Green's function of the time-dependent radiative transfer equation. Biomed Opt Express 3(3):543–551

Lin B et al (1998) Estimation of water cloud properties from satellite microwave, infrared and visible measurements in oceanic environments: 2. Results. J Geophys Res 103:3887

Ma SK (1985) Statistical physics. World Scientific, Philadelphia

Nakajima T, King MD (1990) Determination of the optical thickness and effective particle radius of clouds from reflected solar radiation measurements. Part I: theory. J Atmos Sci 47 (15):1878–1893

Niedre MJ et al (2008) Early photon tomography allows fluorescence detection of lung carcinomas and disease progression in mice in vivo. PNAS 105(49):19126–19131

Patwardhan SV et al (2005) Monte Carlo simulation of light-tissue interaction: three-dimensional simulation for trans-illumination-based imaging of skin lesions. IEEE Trans Biomed Eng 52 (7):1227–1236

Peraiah A (2002) An introduction to radiative transfer: methods and applications in astrophysics. Cambridge University Press, Cambridge

Rybicki GB (1996) Radiative transfer. J Astrophys Astron 17(3–4):95–112
Sawicki J, Kastor N, Xu M (2008) Electric field Monte Carlo simulation of coherent backscattering of polarized light by a turbid medium containing Mie scatterers. Opt Express 16(8):5728–5738
Wang X, Liou K, Ou S (2009) Remote sensing of cirrus cloud vertical size profile using MODIS data. J Geophys Res Atmos 114(D9):D09205
Wehrse R, Kalkofen W (2006) Advances in radiative transfer. Astron Astrophys Rev 13(1–2):3–29
Xu M (2004) Electric field Monte Carlo simulation of polarized light propagation in turbid media. Opt Express 12(26):6530–6539
Xu M, Alfano RR (2005a) Circular polarization memory of light. Phys Rev E 72(6):065601
Xu M, Alfano RR (2005b) Random walk of polarized light in turbid media. Phys Rev Lett 95 (21):213901
Xu M et al (2002) Photon migration in turbid media using a cumulant approximation to radiative transfer. Phys Rev E 65(6):066609
Xu M et al (2001) Photon-transport forward model for imaging in turbid media. Opt Lett 26 (14):1066–1068
Xu M, Cai W, Alfano RR (2004) Multiple passages of light through an absorption inhomogeneity in optical imaging of turbid media. Opt Lett 29(15):1757–1759

Radiative Transfer in Spherically and Cylindrically Symmetric Media

Alexander Kolesov and Nataliya Kropacheva

The theory of light scattering and radiative transfer in various media is successfully used to solve scientific and practical problems of astrophysics, geophysics, oceanology, and other fields of science that study the interaction of the electromagnetic radiation with the matter. Methods of this theory are also used to study the propagation of particles in a substance, in particular, in the study of the diffusion of neutrons. These methods are also applied in studying heat transfer processes.

The theoretical basis for radiation transfer in the plane-parallel media was set forth, for example, by Kourganoff (1952), Davison (1957b), Ambartsumyan (1960, 1998, Chandrasekhar (1960), Busbridge (1960), Sobolev (1963, 1975), Case and Zweifel (1967), Ivanov (1973), van de Hulst (1980), Minin (1988)). In their numerous works, Germogenova, Maslennikov, Marchuk, Bellman, Kalaba, Ueno, Rogovtsov, and other authors presented important results obtained in this theory. Ivanov (1991, 1994) provided a brief history of the theory of radiative transfer.

In many cases, objects with spherical symmetry can be modeled using a plane-parallel media, neglecting by their curvature. In the case of isotropic scattering, this model can be also used to study the light regime in the spherical object when the curvature of the layers cannot be neglected. As a result, the problem of multiple isotropic scattering in a homogeneous sphere and a spherical shell is reduced to the corresponding problem for plane-parallel media.

However, the neglect of the anisotropy of the scattering is not justified to study such astrophysical objects as dust nebulae, planetary atmospheres, and extended shells of stars. The theory of radiative transfer in the spherical shell media should be used in these cases.

A. Kolesov (✉) · N. Kropacheva
Department of General Mathematics and Informatics, St. Petersburg State University, Universitetsky Prospekt, 28, Peterhof, 198504 St. Petersburg, Russia
e-mail: natakr4@gmail.com

© Springer-Verlag Berlin Heidelberg 2016

A. Kokhanovsky (ed.), *Light Scattering Reviews, Volume 11*,
Springer Praxis Books, DOI 10.1007/978-3-662-49538-4_4

Some astrophysical objects (sunspots, accretion disks, coronal rays and others) have axial symmetry. The model of a medium with cylindrical symmetry can be used to study them. This model is also applied in the theory of neutron transport.

In the present review article, the authors focus on studying the radiative transfer in spherically symmetric media. Nevertheless, when it is necessary, we provide an information from the theory of radiative transfer in plane-parallel media. In addition, we will provide a brief overview of the theory of radiative transfer in media with axial (cylindrical) symmetry.

1 Radiation Fields in Infinite Media Illuminated by a Point Source

The problem of a point light source in an infinite homogeneous absorbing and scattering medium plays an important role in the theory of radiative transfer in media with spherical symmetry. The solution of this problem gives the Green's function which allows to determine the radiation field for any spherically symmetric distribution of sources. Knowledge of the light regime in an infinite medium gives an opportunity to calculate radiation fields in other media with spherical symmetry: in a sphere, in an infinite medium with a spherical cavity, and in a spherical envelope. First, the light regime in such media at great distances from the boundary surfaces asymptotically approaches to the light regime in an infinite medium. Second, relations between radiation intensities in these media and in an infinite medium can be used for finding the radiation fields in such media.

The solution of the problem of a point light source in an infinite medium is also used for calculations of thermal neutron fields in spherical nuclear reactors (see Bell and Glasstone 1970).

1.1 Infinite Homogeneous Media with a Planar Source

Let us examine the solution of the problem of the light regime in an infinite homogeneous absorbing and anisotropically scattering medium illuminated by a planar source. This problem is of interest in the sense that its solution is related by some integral equation (see Case et al. 1953; Case and Zweifel 1967) with the problem of the radiative transfer in an infinite medium illuminated by a point source. This circumstance allows finding the exact analytical solution of the problem of a point source, if you know the proper solution of the problem of a planar source.

Case (1960) and Mika (1961) in the cases of isotropic and anisotropic scattering solved the problem of a planar source in an infinite medium by the eigenfunction

method. Case and Zweifel (1967) described the solution of this problem and its results in the book.

The optical properties of the considered homogeneous medium are characterized by the absorption coefficient α, the single scattering albedo λ, and the phase function $x(\cos\gamma)$, where γ is the scattering angle. Let us assume that the phase function is represented as a sum

$$x(\cos\gamma) = \sum_{n=0}^{N} x_n P_n(\cos\gamma), \tag{1}$$

where $P_n(\cos\gamma)$ are Legendre polynomials. The x_n coefficients are the parameters that satisfy the condition $x_0 = 1$ and $|x_n| < 2n + 1$ for any $n \geq 1$.

In the works of Rogovtsov and Borovik (2009), Rogovtsov (2015a, b) the theory was developed for the cases of the phase function square integrable on close interval [−1,1] and satisfying the Hölder condition on this interval.

The functions describing the light regime in an infinite medium with an isotropic point source do not depend on the azimuth. Therefore, considering the problem of a planar source, we limit ourselves to build the Green's function averaged over the azimuth.

Instead of geometric distances z of points of the medium from some arbitrary initial plane, which is coplanar to the planar source, we introduce appropriate optical distances $t = \alpha z (-\infty < t < +\infty)$. The location of the source and the direction of its radiation will be characterized by the coordinate $t_1 (-\infty < t_1 < +\infty)$ and the polar angle $\arccos\eta_1$, respectively. The direction of propagation of the radiation on the optical distance t from the initial plane is determined by the polar angle $\arccos\eta$. We denote that the Green's function averaged over the azimuth by $\overline{G}_\infty(t, \eta; t_1, \eta_1)$.

This function is determined by the equation

$$\eta \frac{\partial \bar{G}_\infty(t, \eta; t_1, \eta_1)}{\partial t} + \bar{G}_\infty(t, \eta; t_1, \eta_1) - \frac{\lambda}{2} \int_{-1}^{1} p(\eta, \eta') \bar{G}_\infty(t, \eta'; t_1, \eta_1) \mathrm{d}\eta' = \frac{1}{2\pi} \delta(t - t_1) \delta(\eta - \eta_1) \tag{2}$$

as well as by the condition of reciprocity (see Case 1960)

$$\bar{G}_\infty(t, \eta; t_1, \eta_1) = \bar{G}_\infty(t_1, -\eta_1; t, -\eta) \tag{3}$$

and by the condition at infinity

$$\lim_{|t| \to \infty} \bar{G}_\infty(t, \eta; t_1, \eta_1) = 0 \tag{4}$$

In Eq. (2) the quantity $p(\eta, \eta')$ is the phase function averaged over the azimuth and $\delta(x)$ is the delta-function.

The Green's function has a jump on the emitting surface. The value of this jump is given by the expression

$$\bar{G}_\infty(t_1+0,\eta;t_1,\eta_1)-\bar{G}_\infty(t_1-0,\eta;t_1,\eta_1)=\frac{\delta(\eta-\eta_1)}{2\pi\eta} \tag{5}$$

The solution of boundary-value problems (2) and (4) is sought in space of linear functionals (see Vladimirov 1976). The mathematical theory of radiative transfer is discussed, for example, by Ershov and Shikhov (1985), Germogenova (1986), Shikhov and Troyanovski (1983), and Vladimirov (1963).

Case (1960) proposed to represent the Green's function at $t < t_1$ and at $t > t_1$ as a superposition of the eigenfunctions (the elementary solutions) $\psi(t, \eta, \nu)$ of the homogeneous radiative transfer equation, i.e., in the form

$$\begin{aligned}\bar{G}_\infty(t,\eta;t_1,\eta_1)&=\mathbb{S}_{\nu>0}c(t_1,\eta_1,\nu)\psi(t,\eta,\nu)\quad \text{for } t>t_1,\\ \bar{G}_\infty(t,\eta;t_1,\eta_1)&=-\mathbb{S}_{\nu>0}c(t_1,\eta_1,-\nu)\psi(t,\eta,-\nu)\quad \text{for } t<t_1.\end{aligned} \tag{6}$$

Here the symbol $\mathbb{S}_{\nu>0}$ denotes summation over all M positive eigenfunctions of the discrete spectrum and integration over the positive eigenfunctions of the continuous spectrum, i.e.,

$$\mathbb{S}_{\nu>0}f(\nu)=\sum_{j=1}^{M}f(\nu_j)+\int_0^1 f(\nu)\,\mathrm{d}\nu. \tag{7}$$

The coefficients $c(t_1, \eta_1, \pm\nu)$ are found from conditions (3), (4), and (5).

The eigenfunctions $\psi(t, \eta, \nu)$ are nontrivial solutions of the equation (see Case and Zweifel 1967)

$$\eta\frac{\partial\psi(t,\eta,\nu)}{\partial t}+\psi(t,\eta,\nu)-\frac{\lambda}{2}\int_{-1}^{1}p(\eta,\eta')\psi(t,\eta',\nu)\mathrm{d}\eta'=0. \tag{8}$$

The solution of this equation has the form

$$\psi(t,\eta,\nu)=R(\eta,\nu)\mathrm{e}^{-\frac{t}{\nu}}. \tag{9}$$

The function $R(\eta, \nu)$ is represented in the form of an expansion in Legendre polynomials $P_n(\eta)$

$$R(\eta, \nu) = \frac{1}{2}\sum_{n=0}^{\infty}(2n+1)R_n(\nu)P_n(\eta). \tag{10}$$

Here quantities $R_n(\nu)$ may be determined from a recursion relation

$$\begin{gathered}(n+1)R_{n+1}(\nu) + nR_{n-1}(\nu) = (2n+1-\lambda x_n)\nu R_n(\nu), \\ R_0(\nu) = 1,\ R_1(\nu) = (1-\lambda)\nu.\end{gathered} \tag{11}$$

Series (10) diverges if $|\nu| \leq 1$, i.e., the value of $R(\eta, \nu)$ is a distribution. Therefore, series (10) cannot be used for numerical calculations, but it can be used to find the angular momenta of this function.

The spectrum of eigenvalues ν of Eq. (2) is mixed. The interval [−1,1] is a continuous range of these values. Discrete values $\nu = \nu_j \quad (j = 1, 2, \ldots, M)$ are obtained by solving the characteristic equation

$$T(\nu) = 1 - \nu\int_{-1}^{1}\frac{\Psi(\eta)}{\nu-\eta}\mathrm{d}\eta = 0, \tag{12}$$

where

$$\Psi(\eta) = \frac{\lambda}{2}\sum_{n=0}^{N}x_n R_n(\eta)P_n(\eta). \tag{13}$$

The spectrum of discrete values of ν was researched by Case (1960), Mika (1961), Maslennikov (1969), Germogenova (1972, 1974, 1978), and Inönü (1970). In the case of isotropic scattering, Eq. (12) has two roots $\pm\nu_1$. When $\lambda = 1$, these roots are infinite. In the case of anisotropic scattering, the number $2M$ of the roots of this equation increases with increasing of the phase function elongation and corresponding increasing of a number $N + 1$ of terms in expression (1). A rigorous presentation of the qualitative and constructive theory of the radiative transfer characteristic equation is given by Rogovtsov and Borovik (2009) and Rogovtsov (2015a, b).

Along with eigenfunctions $\psi(t, \eta, \nu)$, we must introduce corresponding conjugate eigenfunctions

$$\psi^*(t, \eta, \nu) = \psi(t, -\eta, -\nu) = R(\eta, \nu)\mathrm{e}^{\frac{t}{\nu}}. \tag{14}$$

The system of eigenfunctions has the property of orthogonality

$$\int_{-1}^{1}\eta\psi(t, \eta, \nu)\psi^*(t, \eta, \zeta)\mathrm{d}\eta = N(\nu)\delta(\eta, \zeta), \tag{15}$$

where $\delta(\eta,\zeta)=\delta(\eta-\zeta)$ is the delta-function when $v\in[-1,1]$ and $\zeta\in[-1,1]$; $\delta(\eta,\zeta)=\delta_{jk}$ is the Kronecker symbol when $v=v_j,\zeta=\zeta_k$. Here v_j and ζ_k are discrete roots of the characteristic equation (12). In the bounds of the Case method, normalization integrals are given by expressions

$$\begin{aligned} N(v)&=\left\{[T(v)]^2+[\pi v\Psi(v)]^2\right\}v \quad \text{for } |v|\le 1,\\ N(v)&=2v\Psi(v)\left[2v^4\int_0^1\frac{\Psi(\eta)}{(v^2-\eta^2)^2}\mathrm{d}\eta-1\right] \quad \text{for } |v|>1. \end{aligned} \tag{16}$$

Using conditions (4), (5), and (15), Case (1960) and Mika (1961) obtained the following expressions for the Green's function:

$$\bar{G}_\infty(t,\eta;t_1,\eta_1)=\frac{1}{2\pi}\mathbb{S}_{v>0}\frac{R(\eta_1,v)R(\eta,v)}{N(v)}\mathrm{e}^{-\frac{t-t_1}{v}} \quad \text{for } t>t_1, \tag{17}$$

$$\bar{G}_\infty(t,\eta;t_1,\eta_1)=\frac{1}{2\pi}\mathbb{S}_{v>0}\frac{R(-\eta_1,v)R(-\eta,v)}{N(v)}\mathrm{e}^{-\frac{t_1-t}{v}} \quad \text{for } t<t_1. \tag{18}$$

It is usually assumed that $t_1 = 0$.

In the case of a planar isotropic source, the Green's function is given by the expression

$$\bar{G}_\infty(t,\eta;0)=\frac{1}{2}\int_{-1}^{1}\bar{G}_\infty(t,\eta;0,\eta_1)\mathrm{d}\eta_1. \tag{19}$$

Then, expressions (17) and (18) take the following form

$$\bar{G}_\infty(t,\eta;0)=\frac{1}{4\pi}\mathbb{S}_{v>0}\frac{R(\eta,v)}{N(v)}\mathrm{e}^{-\frac{t}{v}} \quad \text{for } t>0, \tag{20}$$

$$\bar{G}_\infty(t,\eta;0)=\frac{1}{4\pi}\mathbb{S}_{v>0}\frac{R(-\eta,v)}{N(v)}\mathrm{e}^{\frac{t}{v}} \quad \text{for } t<0. \tag{21}$$

A planar isotropic source can be represented as a set of point isotropic sources of luminosity L, uniformly distributed with a surface density l. Thus, the quantity $l\cdot L\cdot\bar{G}_\infty(t,\eta;0)$ is the full intensity of the radiation, i.e., the sum of intensities of the diffuse radiation and the radiation coming to this point of the medium directly from the source. Therefore, in this case the source function can be represented in the form

$$B_\infty(t,\eta) = \frac{\lambda l L}{2} \int_{-1}^{1} \bar{G}_\infty(t,\eta';0) p(\eta,\eta') \mathrm{d}\eta', \tag{22}$$

From Eqs. (1), (10), (20), (21), and (22), we obtain the following expressions for this function

$$\begin{aligned} B_\infty(t,\eta) &= \sum_{n=0}^{N} x_n B_n^\infty(t) P_n(\eta) \quad \text{for } t > 0, \\ B_\infty(t,\eta) &= \sum_{n=0}^{N} (-1)^n x_n B_n^\infty(t) P_n(\eta) \quad \text{for } t < 0, \end{aligned} \tag{23}$$

where

$$B_n^\infty(t) = \frac{\lambda l L}{8\pi} \mathbb{S}_{v>0} \frac{R(\eta,v)}{N(v)} \mathrm{e}^{-\frac{|t|}{v}}. \tag{24}$$

These expressions take place for the case of polynomial phase functions.

1.2 The Relationship Between Radiation Field Characteristics in an Infinite Medium with Point or Plane Sources

We consider the problem of light scattering in an infinite homogeneous medium illuminated by an isotropic point sources of luminosity L. We denote the intensity of the diffuse radiation travelling in a direction characterized by a polar angle arccos μ at an optical distance τ from the point source by $I_\infty(\tau,\mu)$ and the corresponding source function by $S_\infty(\tau,\mu)$. These functions are defined by the radiative transfer equation

$$\mu \frac{\partial I_\infty(\tau,\mu)}{\partial \tau} + \frac{1-\mu^2}{\tau} \frac{\partial I_\infty(\tau,\mu)}{\partial \mu} + I_\infty(\tau,\mu) = S_\infty(\tau,\mu) \tag{25}$$

and the radiative equilibrium equation

$$S_\infty(\tau,\mu) = \frac{\lambda}{2} \int_{-1}^{1} p(\mu,\mu') I_\infty(\tau,\mu') \mathrm{d}\mu' + \frac{\lambda L \alpha^2}{16\pi^2} x(\mu) \frac{\mathrm{e}^{-\tau}}{\tau^2}. \tag{26}$$

The function $I_\infty(\tau,\mu)$ must satisfy the condition of its attenuation at large distances from the source, i.e.,

$$\lim_{\tau\to\infty} I_\infty(\tau,\mu) = 0. \tag{27}$$

Solving boundary-value problem (25), (26) is reduced to finding the source function $S_\infty(\tau,\mu)$. Knowing this function, one can determine the radiation intensity $I_\infty(\tau,\mu)$. From Eq. (25) with boundary condition (27), we find

$$I_\infty(\tau,\mu) = \int_0^\infty S_\infty\left(\sqrt{\tau^2 - 2\tau\mu s + s^2}, \frac{\mu}{\sqrt{\tau^2 - 2\tau\mu s + s^2}}\right) \mathrm{e}^{-s}\mathrm{d}s. \tag{28}$$

This expression is suitable for calculations.

We shall write the integral relationship between source functions $B_\infty(t,\eta)$ and $S_\infty(\tau,\mu)$. The function $S_\infty(\tau,\mu)$ can be represented as a sum

$$S_\infty(\tau,\mu) = \sum_{n=0}^{N} x_n S_n^\infty(\tau) P_n(\mu), \tag{29}$$

similar to representation (23) of the function $B_\infty(t,\eta)$. In Case et al. (1953) it was shown that

$$B_n^\infty(t) = \frac{2\pi l}{\alpha^2} \int_{|t|}^\infty S_n^\infty(\tau) P_n\left(\frac{t}{\tau}\right) \tau \,\mathrm{d}\tau. \tag{30}$$

Using the Mellin integral transform from Eq. (30), we obtain the following expressions for the quantities $S_n^\infty(\tau)$ (see Kolesov 1983):

$$S_n^\infty(\tau) = -\frac{\alpha^2}{2\pi l \tau^2}\left[\tau \frac{\mathrm{d}B_n^\infty(\tau)}{\mathrm{d}\tau} - \frac{n(n+1)}{2} B_n^\infty(\tau) - \sum_{m=0}^{m_n} \frac{(-1)^m}{(2m)!!} \frac{(2n-2m-1)!!}{(n-2m-2)!} \tau^{-n+2m+1} \int_\tau^\infty t^{n-m-2} B_n^\infty(t)\mathrm{d}t\right], \tag{31}$$

where $m_n = \frac{1}{2}(n - 2 - \beta_n)$,

$$\beta_0 = 1, \quad \beta_n = \sum_{m=0}^{n}\left(1 - \frac{x_m}{2m+1}\right) \quad \text{for } n \geq 1. \tag{32}$$

The sum in formula (31) is equal to zero if $m_n < 0$.

Dolin and Luchinin (1971) obtained expression (31) by another way.

Thus, the solution of the problem of a point source in an anisotropically scattering and absorbing infinite medium is expressed through the solution of the corresponding problem of a planar source in the same medium.

1.3 Exact Expressions for the Source Function and the Radiation Intensity in an Infinite Medium Illuminated by an Isotropic Point Source

Substitution of expression (24) into Eq. (31) leads to the following exact formula for the quantities $S_n^{\infty}(\tau)$:

$$S_n^{\infty}(\tau) = \frac{\lambda L \alpha^2}{16\pi^2} \mathbb{S}_{\nu>0} \frac{f_n(\tau, \nu)}{N(\nu)}, \tag{33}$$

where

$$f_n(\tau, \nu) = \frac{R_n(\nu)}{\tau\nu} \sqrt{\frac{2\tau}{\pi\nu}}\, K_{n+\frac{1}{2}}\left(\frac{\tau}{\nu}\right). \tag{34}$$

Here $K_{n+\frac{1}{2}}(z)$ is the modified Bessel function of the 3rd kind with a half integer index (see Abramowitz and Stegun 1964).

Thus, the source function is given by the following exact expression

$$S_{\infty}(\tau, \mu) = \frac{\lambda L \alpha^2}{16\pi^2} \sum_{n=0}^{N} x_n P_n(\mu) \mathbb{S}_{\nu>0} \frac{f_n(\tau, \nu)}{N(\nu)}. \tag{35}$$

Expression (35) for the source function takes into account both the diffuse radiation and the direct radiation received at a given point of the medium without scattering.

Representing the intensity of the diffuse radiation $I_{\infty}(\tau, \mu)$ in the form of the expansion in Legendre polynomials and using Eq. (33), we find the expression for the function $I_{\infty}(\tau, \mu)$:

$$I_{\infty}(\tau, \mu) = \frac{L\alpha^2}{16\pi^2} \sum_{n=0}^{\infty} (2n+1)\, P_n(\mu) \left[\mathbb{S}_{\nu>0} \frac{f_n(\tau, \nu)}{N(\nu)} - \frac{\mathrm{e}^{-\tau}}{\tau^2} \right]. \tag{36}$$

It must be emphasized that expression (36) contains a series but not a sum of a finite number of Legendre polynomials. This expression gives the representation of the diffuse radiation intensity in the form of a distribution described by a divergent series. For numerical calculations, expression (36) is unusable. However, the angular momenta of the quantity $I_{\infty}(\tau, \mu)$ obtained through this expression are

usual functions. For example, in the same way, we may find expressions for the mean intensity $J_\infty(\tau)$ and the radiation flux $H_\infty(\tau)$:

$$J_\infty(\tau) = \frac{1}{2}\int_{-1}^{1} I_\infty(\tau,\mu)\mathrm{d}\mu = \frac{L\alpha^2}{16\pi^2\tau}\left[\mathbb{S}_{\nu>0}\frac{\mathrm{e}^{-\frac{\tau}{\nu}}}{\nu N(\nu)} - \frac{e^{-\tau}}{\tau}\right], \tag{37}$$

$$H_\infty(\tau) = 2\pi\int_{-1}^{1} I_\infty(\tau,\mu)\mu\mathrm{d}\mu = \frac{L\alpha^2}{4\pi\tau^2}\left[(1-\lambda)\mathbb{S}_{\nu>0}\frac{\tau+\nu}{N(\nu)} - \mathrm{e}^{-\tau}\right]. \tag{38}$$

These expressions are suitable for numerical calculations in the cases of sufficiently simple phase functions.

The exact expression for $I_\infty(\tau,\mu)$, which can be used for the calculations of the radiation field in an infinite medium, is obtained by substituting expression (35) into (28).

This expression has the following form:

$$I_\infty(\tau,\mu) = \frac{\lambda L\alpha^2}{16\pi^2}\sum_{n=0}^{N} x_n \int_0^\infty \mathbb{S}_{\nu>0}\frac{f_n\left(\sqrt{\tau^2-2\tau\mu s+s^2},\nu\right)}{N(\nu)} P_n\left(\frac{\tau\mu-s}{\sqrt{\tau^2-2\tau\mu s+s^2}}\right)\mathrm{e}^{-s}\,\mathrm{d}s. \tag{39}$$

Kolesov and Perov (1987) used this expression for calculations of the radiation field in an infinite dust nebula illuminated by a star (see Appendix A).

2 The Radiation Field in an Infinite Medium with a Spherical Symmetric Distribution of Sources

The Case method of the solution of the radiative transfer equation, designed for media of planar geometry, was generalized for the case of media with spherical symmetry.

The theory of multiple scattering of light in an anisotropically scattering and absorbing infinite medium with an arbitrary spherically symmetric distribution of radiation sources was developed by Kolesov (1984a, 1994) using this method. Let us consider this theory.

2.1 The System of Eigenfunctions of the Homogeneous Equation of Radiative Transfer in a Spherically Symmetric Medium

In Sect. 1 of the present paper, we considered the problem of the light regime in an infinite medium illuminated by a point source. Here we consider an infinite homogeneous medium with an arbitrary spherically symmetric source distribution.

The eigenfunctions of the radiative transfer equation in this medium are nonzero solutions of the homogeneous equation

$$\mu\frac{\partial f(\tau,\mu,\nu)}{\partial\tau}+\frac{1-\mu^2}{\tau}\frac{\partial f(\tau,\mu,\nu)}{\partial\mu}+f(\tau,\mu,\nu)-\frac{\lambda}{2}\int_{-1}^{1}p(\mu,\mu')f(\tau,\mu',\nu)\mathrm{d}\mu'=0. \tag{40}$$

Partial solutions $\tilde{f}(\tau,\mu,\nu)$ of this equation may be reprinted in the form of the expansion in Legendre polynomials, i.e.,

$$\tilde{f}(\tau,\mu,\nu)=\frac{1}{\nu\sqrt{2\pi\tau\nu}}\sum_{n=0}^{\infty}(2n+1)\,P_n(\mu)R_n(\nu)K_{n+\frac{1}{2}}\left(\frac{\tau}{\nu}\right). \tag{41}$$

The coefficients of such expansion are given by expression (34).

We note that the existence of these partial solutions of the homogeneous transport equation was already known in the 1940s (see Davison 1957b). In particular, they were obtained by solving the homogeneous transport equation in a spherically symmetric medium under the assumption of the isotropy of scattering by Laletin (1966, 1974a) and Nonnenmacher (1967).

When solving radiative transfer problems in an infinite medium, we are usually interested in solutions that are regular at the origin $\tau=0$ and tends to zero at infinity $(\tau\to\infty)$. We have to use the eigenfunctions of the homogeneous transport equation, which provide the possibility of constructing such solutions.

When $\nu>0$, functions (41) satisfy the specified condition at infinity, but they are not regular at the origin. The eigenfunctions $f(\tau,\mu,\nu)=\tilde{f}(\tau,\mu,\nu)+\tilde{f}(\tau,\mu,-\nu)$ are regular at $\tau=0$. These functions have the form

$$f(\tau,\mu,\nu)=\frac{\sqrt{\pi}}{\nu\sqrt{2\tau\nu}}\sum_{n=0}^{\infty}(-1)^{n+1}(2n+1)P_n(\mu)R_n(\nu)I_{n+\frac{1}{2}}\left(\frac{\tau}{\nu}\right),\quad(\nu<0), \tag{42}$$

where $I_{n+\frac{1}{2}}(z)$ is the modified Bessel function of the 1st kind.

Thus, when solving radiative transfer problems, it is more convenient to use the following eigenfunctions

$$\begin{aligned} f(\tau,\mu,\nu) &= \tilde{f}(\tau,\mu,\nu) \quad (\nu>0),\\ f(\tau,\mu,\nu) &= \tilde{f}(\tau,\mu,\nu)+\tilde{f}(\tau,\mu,-\nu) \quad (\nu<0) \end{aligned} \tag{43}$$

instead of the eigenfunctions $\tilde{f}(\tau,\mu,\nu)$. The eigenfunctions

$$\begin{aligned} f^*(\tau,\mu,\nu) &= f(\tau,-\mu,-\nu) = \tilde{f}(\tau,-\mu,\nu)+\tilde{f}(\tau,-\mu,-\nu) \quad (\nu>0),\\ f^*(\tau,\mu,\nu) &= f(\tau,-\mu,-\nu) = \tilde{f}(\tau,-\mu,-\nu) \quad (\nu<0) \end{aligned} \tag{44}$$

are used as the conjugate eigenfunctions.

The eigenfunctions $f(\tau,\mu,\nu)$ and $f^*(\tau,\mu,\nu)$ have the property of orthogonality:

$$\int_{-1}^{1} f(\tau,\mu,\nu)f^*(\tau,\mu,\zeta)\mu\,\mathrm{d}\mu = -\frac{N(\tau)}{\tau^2\nu^2}\delta(\nu,\zeta). \tag{45}$$

Let us consider the calculation of the eigenfunctions. The function $f(\tau,\mu,\nu)$ with $\nu<0$ and $f^*(\tau,\mu,\nu)$ with $\nu>0$ are represented by expansions in a convergent series in Legendre polynomials. Therefore, these expansions can be used in calculations.

The functions $f(\tau,\mu,\nu)$ with $\nu>0$ and $f^*(\tau,\mu,\nu)$ with $\nu<0$ are distributions described by divergent series. The angular momenta of these quantities are usual functions and can be used in the calculations. To calculate these eigenfunctions it is necessary to find a formal solution of Eqs. (40), analogous to the solution (28) of Eq. (25), i.e.

$$\begin{aligned} f(\tau,\mu,\nu) &= \frac{\lambda}{2}\sum_{n=0}^{N} x_n \int_0^{\infty} f_n\left(\sqrt{\tau^2-2\tau\mu s+s^2},\nu\right)\\ &\quad\times P_n\left(\frac{\tau\mu-s}{\sqrt{\tau^2-2\tau\mu s+s^2}}\right)\mathrm{e}^{-s}\mathrm{d}s \quad (\nu>0), \end{aligned} \tag{46}$$

$$\begin{aligned} &f^*(\tau,\mu,\nu)\\ &\quad= \frac{\lambda}{2}\sum_{n=0}^{N}(-1)^{n+1}x_n \int_0^{\infty} f_n\left(\sqrt{\tau^2+2\tau\mu s+s^2},\nu\right)\\ &\qquad\times P_n\left(\frac{\tau\mu+s}{\sqrt{\tau^2+2\tau\mu s+s^2}}\right)\mathrm{e}^{-s}\mathrm{d}s \quad (\nu<0). \end{aligned} \tag{47}$$

In the case of pure scattering when $\lambda = 1$, the characteristic equation (12) has infinite roots $\pm\nu_1 = \pm\infty$ ($k = 0$). Then the corresponding eigenfunctions take the following form:

$$f(\tau, \mu, +\infty) = \frac{1}{2}\sum_{n=0}^{\infty} \frac{n!}{\beta_n} \cdot \frac{P_n(\mu)}{\tau^{n+1}}, \tag{48}$$

$$f(\tau, \mu, -\infty) = 1, \tag{49}$$

$$f^*(\tau, \mu, +\infty) = -1, \tag{50}$$

$$f^*(\tau, \mu, -\infty) = \frac{1}{2}\sum_{n=0}^{\infty} (-1)^{n+1} \frac{n!}{\beta_n} \cdot \frac{P_n(\mu)}{\tau^{n+1}}. \tag{51}$$

In formulae (48) and (51), quantities β_n are given by expression (32).

2.2 Green's Function for the Radiative Transfer Equation in an Infinite Medium with a Spherically Symmetric Distribution of the Sources

Let $G_\infty(\tau, \mu; \tau_1, \mu_1)$ be the Green's function for an infinite medium with a conical sources located on a spherical surface $\tau = \tau_1$ and emitting at angle $\arccos \mu_1 (-1 \le \mu_1 \le 1)$ to the radius vector. This function satisfies the equation

$$\begin{aligned} &\mu \frac{\partial G_\infty(\tau, \mu; \tau_1, \mu_1)}{\partial \tau} + \frac{1-\mu^2}{\tau} \frac{\partial G_\infty(\tau, \mu; \tau_1, \mu_1)}{\partial \mu} + G_\infty(\tau, \mu; \tau_1, \mu_1) \\ &\quad - \frac{\lambda}{2} \int_{-1}^{1} p(\mu, \mu') G_\infty(\tau, \mu'; \tau_1, \mu_1) \mathrm{d}\mu' \\ &\quad = \frac{1}{2\pi\tau^2} \delta(\tau - \tau_1) \delta(\mu - \mu_1), \end{aligned} \tag{52}$$

as well as the condition of reciprocity

$$G_\infty(\tau, \mu; \tau_1, \mu_1) = G_\infty(\tau_1, -\mu_1; \tau, -\mu), \tag{53}$$

and conditions of the regularity at the center of symmetry (when $\tau = 0$) and of the trending to zero at infinity ($\tau \to \infty$).

Equation (52) for the Green's function becomes homogeneous when $\tau \neq \tau_1$. Therefore, this function is represented as a linear combination of the eigenfunctions $f(\tau, \mu, \nu)$. When $\tau = \tau_1$, it has a gap, which is determined by expression

$$G_\infty(\tau_1 + 0, \mu; \tau_1, \mu_1) - G_\infty(\tau_1 - 0, \mu; \tau_1, \mu_1) = \frac{\delta(\mu - \mu_1)}{2\pi\tau_1^2 \mu}. \tag{54}$$

Using the expansion of this function in the eigenfunctions, as well conditions (45) and (54), we obtain the following expressions:

$$G_\infty(\tau,\mu;\tau_1,\mu_1) = -\frac{1}{2\pi}\mathbb{S}_{\nu>0}\frac{\nu^2}{N(\nu)}f^*(\tau_1,\mu_1,\nu)f(\tau,\mu,\nu) \quad (\tau>\tau_1), \tag{55}$$

$$G_\infty(\tau,\mu;\tau_1,\mu_1) = -\frac{1}{2\pi}\mathbb{S}_{\nu>0}\frac{\nu^2}{N(\nu)}f^*(\tau_1,\mu_1,-\nu)f(\tau,\mu,\nu) \quad (\tau<\tau_1). \tag{56}$$

When the radiation source is isotropic, the integration of expressions (55) and (56) over μ_1 from -1 to 1 leads to the expressions

$$G_\infty(\tau,\mu;\tau_1) = \frac{1}{2\pi\tau_1}\mathbb{S}_{\nu>0}\frac{\nu}{N(\nu)}f(\tau,\mu,\nu)sh\frac{\tau_1}{\nu} \quad (\tau>\tau_1), \tag{57}$$

$$G_\infty(\tau,\mu;\tau_1) = -\frac{1}{4\pi\tau_1}\mathbb{S}_{\nu>0}\frac{\nu}{N(\nu)}f(\tau,\mu,-\nu)\mathrm{e}^{-\frac{\tau_1}{\nu}} \quad (\tau<\tau_1). \tag{58}$$

Trending τ_1 to zero in formula (57), we obtain the expression of the Green's function for an infinite medium illuminated by an isotropic point source located at the origin

$$G_\infty(\tau,\mu;0) = \frac{1}{2\pi}\mathbb{S}_{\nu>0}\frac{f(\tau,\mu,\nu)}{N(\nu)} \tag{59}$$

Meyer and Jacobs (1970) also obtained this expression using another method.

The intensity of the diffuse radiation $I_\infty(\tau,\mu;\tau_1,\mu_1)$ is expressed through the Green's function $G_\infty(\tau,\mu;\tau_1,\mu_1)$ by the relationship

$$I_\infty(\tau,\mu;\tau_1,\mu_1) = \frac{L\alpha^2}{4\pi}G_\infty(\tau,\mu;\tau_1,\mu_1)|\mu_1|. \tag{60}$$

We note that the eigenfunctions used for calculations of the Green's function are not divergent expansions in Legendre polynomials but the formal solutions of Eq. (40). Therefore, the following formulae take the place for $I_\infty(\tau,\mu;\tau_1,\mu_1)$ when $\tau>\tau_1$ or when $\tau<\tau_1$:

$$I_\infty(\tau,\mu;\tau_1,\mu_1) = -\frac{\lambda L\alpha^2|\mu_1|}{16\pi^2}\sum_{n=0}^{N}x_n\mathbb{S}_{\nu>0}\frac{\nu^2}{N(\nu)}f^*(\tau_1,\mu_1,\nu)$$
$$\times\int_0^\infty f_n\left(\sqrt{\tau^2-2\tau\mu s+s^2},\nu\right)P_n\left(\frac{\tau\mu-s}{\sqrt{\tau^2-2\tau\mu s+s^2}}\right)\mathrm{e}^{-s}\mathrm{d}s \quad (\tau>\tau_1), \tag{61}$$

$$I_\infty(\tau,\mu;\tau_1,\mu_1)=\frac{\lambda L\alpha^2|\mu_1|}{16\pi^2}\sum_{n=0}^{N}(-1)^{n+1}x_n\mathbb{S}_{\nu>0}\frac{\nu^2}{N(\nu)}f(\tau,\mu,-\nu)$$
$$\times\int_0^\infty f_n\left(\sqrt{\tau_1^2+2\tau_1\mu_1 s+s^2},\nu\right)P_n\left(\frac{\tau\mu_1+s}{\sqrt{\tau_1^2+2\tau_1\mu_1 s^2+s^2}}\right)\mathrm{e}^{-s}\mathrm{d}s\quad(\tau<\tau_1). \tag{62}$$

3 Radiation Fields in a Sphere and in a Spherical Envelope

3.1 Main Methods for Solving Equation of Radiative Transfer in Spherically Symmetric Media

The earliest works, in which the spherical radiative transfer equation has been used for solution of astrophysical problems, were written in 1920s and 1930s. They were concerned with the problem of the radiative equilibrium in solar and stellar atmospheres. In these papers, the temperature distribution in extended photospheres was determined, the energy distribution in their continuous spectra was found, and the problem of absorption lines formation was discussed. Subsequently, the theory of radiative transfer in spherically symmetric media was used for the planetary atmospheres and the dust nebulae.

Exact, approximate, and numerical methods of solving the problems of the neuron diffusion in media of different geometric shapes began to develop intensively in 1940s in connection with the problem of calculation of nuclear reactors. The neuron diffusion and radiative transfer problems are mathematical equivalent. Therefore, these methods are also applied in calculations of radiation fields in astrophysical objects.

It should be noted that analytical methods based on exact or approximate theory are rather rarely used in astrophysics and physics. It happens, because analytical methods are too cumbersome in practically interesting cases of heterogeneous media, nonmonochromatic light scattering, and complicated spatial and temporal distributions of energy sources. Therefore, various numerical methods of solving the transport equation were developed for practical applications. Both methods of the direct numerical solution of the transport equation and methods based on the results of exact or approximate theory are widely used.

You can find reviews of various methods of solving the transport equation in a number of articles and monographs, for example, in Bell and Glasstone (1970), Greenspan et al. (1968), Davison (1957b), Germogenova (1971, 1986), Marchuk (1961), Marchuk and Lebedev (1981). Crosbie and Khalil (1972) adduced the extensive bibliography of works on the theory of radiative transfer in media with spherical symmetry.

Here, we give a brief overview of the most used methods for solving radiative transfer problems in media with spherical symmetry.

Taylor (1927) and McCrea (1928) studied for the first time the radiation field in spherical shells in the case of isotropic scattering. In the first of these works, the radiative transfer equation was solved in the Schwarzschild–Schuster approximation (see Schwarzschild 1914). In the second work, the two-stream approximation was proposed. In these approximations, the intensity of radiation is averaged in two intervals of variation of the angular variable μ, i.e., $1 \le \mu \le \mu_r$ and $\mu_r < \mu \le 1$. In contrast to the Schwarzschild–Schuster approximation, where $\mu_r = 0$, in the two-beam approximation it is postulated that arccos μ_r is the visible angle size of the inner radius r_* of the shell, if the point of the observation is located at distance r from the center of symmetry, i.e., $\mu_r^2 = 1 - \left({}^{r_*}/_{r}\right)^2$.

Later, Unno and Kondo (1976) proposed a generalized two-stream approximation used in the case of shells of large optical thickness. In this approximation, the value μ_r is determined by solving an auxiliary differential equation. In Unno and Kondo (1977) it was considered an absorbing and anisotropically scattering spherical shell. In Takenti (1979), this method was used to study the radiation field in an infinitely long shell with the volume absorption coefficient, decreasing with distance from the center of symmetry by a power law.

Wilson et al. (1980) developed a three-stream approximation based on the averaging of the radiation intensity in three intervals of variation of the angular variable μ, i.e., $-1 \le \mu < 0$, $0 \le \mu \le \mu_r$, and $\mu_r < \mu \le 1$.

The Eddington approximation (see Eddington, 1926) as well as its numerous modifications are frequently used for solving the transport equation in various problems of astrophysics and physics. This method was first used by Kosirev (1934) and Chandrasekhar (1934a, b, 1935) in the case of isotropically scattering media with spherical symmetry.

The method of spherical harmonics, which is the generalization of the Eddington method, is one of the most effective methods for solving the transport problems (see, for example, Greenspan et al. 1968; Davison 1957b; Marchuk 1961; Vladimirov 1963). This method is based on the expression of the sought radiation intensity (or phase density of neutrons) in spherical harmonics. The approximate expression for the intensity of radiation, called P_N-approximation, is sought in the form of a partial sum of this expansion, consisting of $N + 1$ members. Substitution of this expression in the integro-differential transfer equation leads to a system of differential equations for the expansion coefficients. This system is solved with boundary conditions corresponding to conditions of the considered physical or astrophysical problems. Thus, P_1-approximation coincides with the Eddington approximation.

In the case of media with spherical symmetry, Marshak (1947) first used the method of spherical harmonics. He applied this method to the problem of the distribution of neutrons in an infinite homogeneous medium with a completely black spherical cavity in the case of pure isotropic monochromatic scattering. As applied to a spherical symmetric media, this method is described, in particular, by

Davison (1957a), Bell and Glesston (1970), Marchuk and Lebedev (1981). In the last two books, the case of anisotropic scattering is also reviewed. DeBar (1967) applied this method to the solution of the nonstationary radiative transfer equation.

When using the spherical harmonics method, angular momenta of the radiation intensity are sometimes applied instead of the expansion coefficients of the radiation intensity in Legendre polynomials. Simonneau (1976, 1978, 1980) developed the numerical method for solving the radiative transfer problem in inhomogeneous media of spherical geometry based on the usage of these momenta.

Yvon (1957) proposed the method of dual P_N-approximation, which is a modification of the method of spherical harmonics. In this approximation, the radiation intensity (or the phase density of neutrons) is sought in the form of separate expansions in Legendre polynomials in each of two intervals of the variation of the angular variable, i.e., $-1 \le \mu \le 0$ and $0 \le \mu \le 1$. This method was used for solving radiative transfer problems in a homogeneous sphere by Drawbaugh and Noderer (1959), Wilson and Sen (1964).

Yvon method is close to the method of the angular momenta on the hemisphere, which was proposed by Özisik et al. (1975). The numerical method for solving the transport equation in media of spherical geometry in the cases of monochromatic scattering and scattering with radiation frequency redistribution was developed on the base of this method.

Sobolev (1943) proposed an approximate method for studying the light regime in a planar medium. This method consists of the following. Light scattering of the first order is taken into account exactly. Nevertheless, the Eddington approximation is used and the exact phase function is replaced by a two-term one in the case of scattering of higher order. Later, Sobolev (1972) used this method for investigating the radiation field in a homogeneous spherical dust nebula. Dorschner (1971), who proposed to describe scattering of higher order in P_N-approximation, generalized this method.

Chou and Tien (1968) developed the numerical method of solving the radiative transfer equation in a spherical shell with anisotropic scattering. This method is based on the usage of the Eddington approximation and the presentation of the radiation intensity in the form of a linear function of the angular variable in each of the three intervals of its variation, i.e., $-1 \le \mu < 0$, $0 \le \mu \le \mu_r$ and $\mu_r < \mu \le 1$. This technique was improved by Viik (1974a, b, c).

Pustylnick (1964, 1965) changed the Eddington method supposing the existence of a relationship between the Eddington multipliers dependent on the radial coordinate r, i.e., the relations of K-integral $K(r)$ and flux $H(r)$ to the mean intensity $J(r)$. Hummer and Rybicki (1971) developed the modified Eddington method based on finding the variable multiplier $K(r)/J(r)$ with the help of iterations. It was used in the case of an absorbing and anisotropically scattered medium by Cassinelli and Hummer (1971). As shown in Hummer and Rybicki (1971) and in Rybicki and Hummer (1975), this method allows to take into account the effect of the peak emission intensity in problems of light scattering in the lengthy shells with the density decreasing with the distance r from the center of symmetry. Chapman

(1966) identified this effect. The radiation intensity $I(\tau, \mu)$ has the pronounced maximum at $\mu = 1$, the value of which increases with r. The numerical algorithm description for solving the radiative transfer equation by this method in the case of pure isotropic scattering, is shown by Hummer et al. (1973). The Eddington multiplier $K(r)/J(r)$ was computed in Wilson et al. (1972) for the case when the volume scattering coefficient is inversely proportional to r^2.

Leung (1975) extended this method to the case of an absorbing and scattering spherically symmetric medium with internal energy sources. It is known as the quasidiffusion method.

Kunasz and Hummer (1974) proposed the numerical method of the radiative transfer equation solution in a medium with spherical symmetry, suitable for the calculation of spectral line profiles. Using this method, both Eddington multipliers, i.e., $K(r)/J(r)$ and $H(r)/J(r)$ were calculated by iteration. Yorke (1986) developed the method for the solution of stationary and non-stationary problems of radiative transfer in continuous spectrum with anisotropic scattering based on the use of these multipliers.

The discrete ordinates method is also widely used for solving radiative transfer problems. For case of plane-parallel media, Wick (1943) suggested the original version of this method developed then by Chandrasekhar (1960). This method consists of the replacement of the integral term in the radiative transfer equation by the Gaussian sum. Chandrasekhar (1945) applied this method to the investigation of the radiation field in an inhomogeneous spherical envelope in the case of the pure isotropic monochromatic scattering under the assumption that the volume coefficient of emission decreases with increasing of the distance from the center of symmetry. By this method, Barkov (1960, 1963) investigated the light regime in a homogeneous sphere with a central point source.

Another variant of the method of discrete ordinates is S_n-method developed by Carlson (see Greenspan et al. 1968). In this method, the integral term in the radiative transfer equation is calculated approximately. The interval of variation of the angular variable ($-1 \le \mu \le 1$) is divided into N sections. In every section, the function is replaced by a linear function of this variable. Davydov and Shikhov (1973, 1975) described the application of this method to media of spherical geometry.

Different methods for the numerical solution of the integro-differential radiative transfer equation are used for computing the light regime in inhomogeneous spherically symmetric media. These methods are based on discretization not only the angle but also the radial variable (see, for example, Grant and Hunt Grant and Hunt 1969a, b; Gruschinske and Ueno 1971; Peraiahand Varghese 1984, 1985a, b). While studying the light scattering with the radiation frequency redistribution, the frequency is also exposed to the discretization.

The radiative and neutron transport boundary-value problems are often solved by numerical methods, in particular, by the Monte Carlo method (see, for example, Marchuk and Mikhailov 1967; Marchuk and Lebedev 1981; Ermakov and Mikhailov 1982; Mikhailov 1987). Kurchakov (1970) and Sanford (1973) applied

the Monte Carlo method to the calculation of radiative fields in spherically symmetric media, for example.

When solving problems of the radiative transfer theory, it is often more convenient to use the integral equation for the source function instead of the integro-differential equation for the radiation intensity. Peierls (1939) derived the integral equation for media of arbitrary geometry in the case of isotropic scattering. This equation for planar and spherically symmetric media will be presented below, in the next section of the present article.

The Peierls equation was solved by various numerical methods. The iteration method was used to solve this equation by Kuznetsov (1951) and Viik (1974a, b, c). Kuznetsov (1951) calculated the source function for an inhomogeneous infinitely long shell surrounding a black sphere, assuming a pure scattering. Viik (1974a, b, c) considered the case of a homogeneous spherical layer under various assumptions about the distribution of internal and external sources. In Kho and Sen (1971), the source function was sought in the form of Neumann series. In Dubi and Horowitz (1978), the difference of the source functions for the sphere and for the infinite medium was expanded in the same series. Gruschinske and Ueno (1971) solved the integral radiative transfer equation by means of the method proposed by Bellman (1968) for solving Fredholm equations. Leong and Sen (1971a) solved this equation, representing its kernel as a sum of degenerate and small kernels. In these works, the phase function was considered to be spherical.

Rogovtsov (1981a, b) and Ganapol (2008) obtained various important invariance relations and integral equations generalizing the Peierls equation.

Calculating the source function for inhomogeneous media of spherical geometry, Schmid-Burgk (1973, 1975), Schmid-Burgk and Scholz (1975), broke these media on several spherical layers. In each of these layers, the product of the square of the distance from the center of symmetry on the source function was described by a polynomial depending on the optical depth of the layer.

Let us now consider methods for solving the radiative transfer equation in media with spherical symmetry, based on the results of the exact theory.

A number of authors solved the problems of the light (or neutrons) propagation in a sphere of finite radius and in a spherical shell in the case of isotropic scattering, by reducing them to the appropriate problems for planar media.

Heaslet and Warming (1965) considered the case of the uniform distribution of energy sources in a sphere and expressed the source function through the fundamental function $\Phi(\tau)$ for a planar layer. Sobolev (1958) introduced this function in the theory. The function $\Phi(\tau)$ is associated with resolvent of the integral radiative transfer equation in a planar layer. These authors expressed the luminosity of a sphere through the angular momenta of X- and Y-functions of Chandrasekhar. Gritton and Leonard (1970) deduced an expression for the amount of heat leaving the surface of a sphere. This expression contains the X- and Y-functions. Van de Hulst (1987, 1988, 1994) expressed the double angular momenta of the light reflection coefficient from a sphere through the momenta of these functions.

Sobolev (1972) determined the average number of scattering of a light quantum, radiated at a certain optical depth in a sphere, in the diffusion process until its exit

through the boundary surface or until its true absorption. This value is expressed through the fundamental function $\Phi(\tau)$. Sabashvili (1973) made numerical calculations of the average number of quantum scattering in a sphere. Knowledge of this value allows us to determine the luminosity of a sphere for an arbitrary (not necessarily spherically symmetric) distribution of sources. Sobolev (1972) considered the cases when the sources in a sphere were evenly distributed and when a sphere was illuminated by a point source located at an arbitrary distance from the center. Loskutov (1974) calculated the luminosity of a sphere illuminated by parallel rays.

Case et al. (1970), solving the problem of the neutron distribution in spherically symmetric media with isotropic scattering by the method of Green's function, received an expansion of the neutrons phase density in the eigenfunctions for this problem. For the expansion coefficients, they brought the same linear singular equations, which were obtained in the case of planar media. However, as shown by Sahni (1975) in that way we get the correct solution only for the "internal" problem of the neutron transport. This method does not resolve the "external" problem, i.e., the problem of neutron transport in an infinite medium with a spherical absolutely black cavity.

Leonard and Mullikin (1964) found in the case of isotropic scattering the relation between the resolvents of the integral transport equation for a sphere of optical radius τ_0 and a planar layer of optical thickness $2\tau_0$. Using this relation, Erdman and Siewert (1968) expressed the source function for a sphere through the Case eigenfunctions for a planar layer.

Wilson and Sen (1973a) obtained the expression for the resolvent of the integral radiative transfer equation for a sphere through two auxiliary functions defined by Fredholm integral equations. Ueno (1974) expressed this quantity through a single auxiliary function similar to the fundamental function $\Phi(\tau)$ for a planar layer. Poon and Ueno (1974) investigated analytic properties of this auxiliary function.

Shkurpelov and Ershov (1967), Shkurpelov et al. (1970) transformed the integral radiative transfer equation for a sphere and got integro-differential equation for the auxiliary quantity, which has the meaning of the radiation intensity (or phase density of neutrons) in a planar layer. This integro-differential equation was solved using the Case method. Gorelov and Yuferev (1971), Gorelov et al. (1972), Sheaks (1972), and Sahni (1975) solved this equation in the same way for a sphere consisting of homogeneous layers, for a homogeneous infinite medium with a completely black spherical cavity, and for a homogeneous spherical envelope. When solving transport problems in this way, explicit analytical expressions for the expansion coefficients of the radiation intensity (or phase density neutron) in the Case eigenfunctions does not found. These coefficients are found by numerical solution of linear singular equations.

In Siewert and Grandjean (1979), Siewert and Maiorino (1979), Siewert and Thomas (1984), the auxiliary integro-differential equation has been solved by F_N-method proposed by Siewert and Benoist (1979) for the case of planar media. This method is based on the expansion of the phase density of neutrons (or the radiation intensity) on the boundary surface of a medium in a series in powers of the angular variable and the replacement of this series by the partial sum with $N + 1$ terms.

Laletin (1969b) developed the method of surface pseudosources (G_N-approximation) for solving boundary-value problems of the transfer theory. In this method, the phase density of neutrons (or radiation intensity) is expressed through Green's function for the infinite homogeneous medium, the function of the source distribution in the medium and the distribution function of the so-called pseudosources. These expressions are sought in the form of the expansion in the series of spherical functions. This series is approximately replaced by a partial sum.

Ershov and Shikhov (1977) described methods for solving boundary-value problems of the transfer theory, based on the usage of integral transformations. Solutions of the radiative transfer equations were obtained by means of the transformations of Radon and Fourier (see, for example, Ershov and Shikhov 1972, 1973).

To solve the problems of the reflection and transmission of light by plane-parallel media, Ambartsumyan (1943) proposed the method of layers combination. He used the invariance principle consisting in the following: the intensity of the radiation emerging from a semi-infinite medium does not change when an additional layer with the same optical properties is added to this medium. This principle is performed only in the case of plane-parallel media. However, the method of invariant embedding (see monograph by Casti and Kalaba 1973), which is a generalization of the method of combination of layers, is applicable to the media of nonplanar geometries.

When this method is used, it is taken into account all changes of the intensity of reflected radiation brought about by this layer. The processes of absorption and re-emission of light in an additional thin layer added to the medium cause these changes. The resulting equations contain derivatives with respect to coordinates characterizing the geometric position of the boundary surface of the medium.

Bellman and other authors (Bailey 1964; Bailey and Wing 1964; Bellman and Kalaba 1957; Bellman et al. 1959, 1960, 1964) used the method of invariant embedding for media with a spherical symmetry with isotropic light scattering. Ueno et al. (1971) derived integro-differential equations for the function of the reflection of light from a sphere, as well as for the functions of the reflection and transmission of light by a spherical shell for the case of inhomogeneous anisotropic scattering spherical shells under various conditions on the inner boundary surface. These authors also obtained the integro-differential equation for radiation intensity inside a spherical shell.

Bellman and Kalaba (1965), Bellman et al. (1966) calculated the functions of diffuse reflection of light from a spherical shell. Rybicki (1970) proved that the results of these calculations are wrong. These authors used a system of equations, which is not related to the function of diffuse light reflection. This system is related to the fully light reflection function containing distributions of the delta-function type. This function describes a direct flight of quanta of light through the boundary surface of a shell. As a result, the collateral regular solution of the system of equations has no physical meaning. This solution was mistakenly identified with the function of the diffuse reflection of light.

Rogovtsov (1981a, b, 1983, 1985, 1999) generalized the invariance principle and stated the general invariance principle. Based on this principle, a new variant of the invariance embedding method and the general invariance relation method were developed. This method was used to solve the boundary-value problems of radiative transfer theory for turbid media of various configurations (see Rogovtsov 2010). In particular, Rogovtsov (1986a, b, 1990, 1992a, b) considered radiative fields in spherically and cylindrically symmetric media.

Rogovtsov and Borovik (2009), Rogovtsov et al. (2010) developed and used effective algorithms of reflection function for the case of greatly elongated phase functions.

Sobolev (1975) proposed a probabilistic method for solving radiative transfer problems in plane-parallel media. This method is based on finding the probability that a light quantum originating at a given optical depth in a medium will emerge from the medium in the particular direction. Minin (1961, 1966) generalized this method to the case of anisotropic scattering.

This method was used by Leong and Sen (1968, 1969, 1970, 1972, 1973a, b) in the case of media with spherical symmetry. Leong and Sen (1968) introduced in theory the concept of the probability of light quantum emerging from a spherical shell; they obtained integro-differential equations for functions of the reflection and transmission of light by inhomogeneous spherical shell with anisotropic monochromatic scattering. Wilson and Sen (1973b) introduced function analogues to X and Y functions of Chandrasekhar for inhomogeneous isotropically scattering spherical shell and expressed through these functions the probability of the emergence of a light quantum through external and internal boundary surface of the shell. Nonstationary radiation field in an inhomogeneous spherical shell with isotropic scattering was investigated by the probabilistic method by Leong and Sen (1972) and Leong (1972). The intensity of the radiation emerging from a medium is expressed through auxiliary functions which are defined by integro-differential equations. Using this method, Leong and Sen (1973a) obtained equations for the coefficients of the reflection and transmission of light by a spherical shell in the frequencies of the spectral lines with the complete redistribution of the radiation frequency. Wilson and Sen (1973c) generalized this result to the case of an arbitrary law of the radiation frequency redistribution. Yengibarian (1972) solved the problem of light scattering in a sphere with an arbitrary distribution of sources.

The analytic theory of radiative transfer in moving media is developed in a number of articles. For example, Grachov (1978, 1994) considered the radiative transfer in linearly expanding homogeneous spherical envelope surrounding an opaque sphere. Sen and Wilson (1993) used the generalized Eddington approximation for determining radiation fields in spherically symmetric moving media.

The radiative transfer problem in a planetary atmosphere illuminated by parallel solar radiation was set out by Sobolev (1975), van de Hulst (1980), Minin (1988), Smoktii (1986), Smoktii and Anikonov (2008). This problem is practically important for astrophysics and geophysics.

In Sect. 4 of the present article, we consider in more detail the problem of determining radiation fields in envelopes of limited size on the basis of integral relations between the Green's functions for such media and for an infinite medium.

3.2 The Integral Equation for the Source Function in the Case of a Sphere

Let us examine the Peierls equation in the case of a homogeneous sphere of optical radius τ_0. As in the previous sections of the present article, the optical properties of a medium will be characterized by the volume absorption α, the single scattering albedo λ, and the phase function $x(\cos\gamma)$. The intensity of the diffuse radiation $I(\tau, \mu; \tau_0)$ and the source function $S(\tau, \mu; \tau_0)$ are determined by the equation of radiative transfer

$$\mu \frac{\partial I(\tau, \mu; \tau_0)}{\partial \tau} + \frac{1 - \mu^2}{\tau} \cdot \frac{\partial I(\tau, \mu; \tau_0)}{\partial \mu} + I(\tau, \mu; \tau_0) = S(\tau, \mu; \tau_0) \tag{63}$$

and the equation of radiative equilibrium

$$S(\tau, \mu; \tau_0) = \frac{\lambda}{2} \int_{-1}^{1} p(\mu, \mu') I(\tau, \mu'; \tau_0) \mathrm{d}\mu' + S_0(\tau, \mu; \tau_0), \tag{64}$$

with the boundary condition

$$I(\tau_0, \mu; \tau_0) = 0 \quad (-1 \leq \mu < 0). \tag{65}$$

In Eq. (64), the source term $S_0(\tau, \mu; \tau_0)$ results from scattering of the direct radiation from the source. Condition (65) corresponds to the absence of an external diffuse radiation.

From Eqs. (63) and (64), it is possible to obtain one integral equation that determine the source function. Integrating linear differential Eq. (63) with boundary condition (65), we find

$$I(\tau, \mu; \tau_0) = \int_{\tau}^{\tau_0} S\left[t, -\sqrt{1 - \frac{\tau^2}{t^2}(1 - \mu^2)}; \tau_0\right] \cdot \mathrm{e}^{-\tau\mu - \sqrt{t^2 - \tau^2(1-\mu^2)}} \frac{t\,\mathrm{d}t}{\sqrt{t^2 - \tau^2(1 - \mu^2)}} \quad \text{for } \mu < 0, \tag{66}$$

$$
\begin{aligned}
&I(\tau,\mu;\tau_0)\\
&\quad=\int\limits_{\tau\sqrt{1-\mu^2}}^{\tau} S\left[t,\sqrt{1-\frac{\tau^2}{t^2}(1-\mu^2)};\tau_0\right]\cdot \mathrm{e}^{-\tau\mu+\sqrt{t^2-\tau^2(1-\mu^2)}}\frac{t\,\mathrm{d}t}{\sqrt{t^2-\tau^2(1-\mu^2)}}\\
&\quad+\int\limits_{\tau\sqrt{1-\mu^2}}^{\tau_0} S\left[t,-\sqrt{1-\frac{\tau^2}{t^2}(1-\mu^2)};\tau_0\right]\cdot \mathrm{e}^{-\tau\mu-\sqrt{t^2-\tau^2(1-\mu^2)}}\frac{t\,\mathrm{d}t}{\sqrt{t^2-\tau^2(1-\mu^2)}}\quad \text{for } \mu>0.
\end{aligned}
\tag{67}
$$

By substituting expressions (66) and (67) into Eq. (64), we obtain after transformations the following integral equation for the source function:

$$
\begin{aligned}
\tau S(\tau,\mu;\tau_0)=\frac{\lambda}{2}\int\limits_0^{\tau_0} t\,\mathrm{d}t\int\limits_{|t-\tau|}^{t+\tau} S\left(t,-\frac{y^2-\tau^2+t^2}{2yt};\tau_0\right)\\
\times p\left(\mu,\frac{y^2-t^2+\tau^2}{2y\tau}\right)\mathrm{e}^{-y}\frac{\mathrm{d}y}{y}+\tau S_0(\tau,\mu;\tau_0).
\end{aligned}
\tag{68}
$$

By setting $S(-\tau,\mu;\tau_0)=S(\tau,\mu;\tau_0)$, the Eq. (68) may be rewritten as

$$
\begin{aligned}
\tau S(\tau,\mu;\tau_0)=\frac{\lambda}{2}\int\limits_{-\tau_0}^{\tau_0} t\,\mathrm{d}t\int\limits_{|t-\tau|}^{\infty} S\left(t,-\frac{y^2-\tau^2+t^2}{2y|t|};\tau_0\right)\\
\times p\left(\mu,\frac{y^2-t^2+\tau^2}{2y|\tau|}\right)\mathrm{e}^{-y}\frac{\mathrm{d}y}{y}+\tau S_0(\tau,\mu;\tau_0).
\end{aligned}
\tag{69}
$$

In the case of isotropic scattering, when $p(\mu,\mu')=1$, the source function $S(\tau;\tau_0)$ does not depend on the angular variable μ. Then the Eq. (69) has the form

$$
\tau S(\tau;\tau_0)=\frac{\lambda}{2}\int\limits_{-\tau_0}^{\tau_0} tS(t,\tau_0)\,E_1(|\tau-t|)\mathrm{d}t+\tau S_0(\tau;\tau_0),
\tag{70}
$$

where $E_k(z)$ is the exponential integral, defined by

$$
E_k(z)=\int\limits_1^{\infty}\mathrm{e}^{-zu}\frac{\mathrm{d}u}{u^k}.
\tag{71}
$$

Equation (70) for the quantity $\tau S(\tau;\tau_0)$ in the case of a sphere of optical radius τ_0 has the same form as the Peierls equation in the case of a plane-parallel medium of optical thickness 2 τ_0 (see, for example, Leonard and Mullikin 1964).

Let us examine the case of anisotropic scattering of light, when the phase function $p(\mu, \mu')$ is represented by a sum of a finite number of Legendre polynomials. If the value $S_0(\tau, \mu; \tau_0)$ is also represented by Legendre polynomials, i.e.,

$$S_0(\tau, \mu; \tau_0) = \sum_{n=0}^{N} x_n S_n^0(\tau; \tau_0) P_n(\mu), \tag{72}$$

then the function $S(\tau, \mu; \tau_0)$ has a similar view, i.e.,

$$S(\tau, \mu; \tau_0) = \sum_{n=0}^{N} x_n S_n(\tau; \tau_0) P_n(\mu). \tag{73}$$

The coefficients $S_n(\tau; \tau_0)$ and the intensity of radiation are related by the equation

$$S_n(\tau; \tau_0) = \frac{\lambda}{2} \int_{-1}^{1} P_n(\mu)\, I(\tau, \mu; \tau_0) \mathrm{d}t + S_n^0(\tau; \tau_0). \tag{74}$$

The substitution of expression (73) into Eq. (69) leads us to the following system of linear integral equations for the quantities $S_n(\tau; \tau_0)$:

$$\begin{aligned} \tau S_n(\tau; \tau_0) = {} & \frac{\lambda}{2} \sum_{m=0}^{N} (-1)^m x_n \int_{-\tau_0}^{\tau_0} t S_m(t, \tau_0)\, Q_{mn}(\tau, t) \mathrm{d}t \\ & + \tau S_n^0(\tau; \tau_0), \end{aligned} \tag{75}$$

where

$$Q_{mn}(\tau, t) = \int_{|t-\tau|}^{\infty} P_m\left(\frac{y^2 - \tau^2 + t^2}{2y|t|}\right) P_n\left(\frac{y^2 - t^2 + \tau^2}{2y|\tau|}\right) \mathrm{e}^{-y} \frac{\mathrm{d}y}{y}. \tag{76}$$

We note that the kernel functions $Q_{mn}(\tau, t)$ are symmetric with respect to arguments and indices, e.g., $Q_{mn}(\tau, t) = Q_{nm}(t, \tau)$. They can be expressed through the exponential functions and the exponential integrals.

In the case of isotropic scattering, the source function is determined by Eq. (70), the kernel of which depends on the modulus of the difference of the arguments. Such situation is typical for problems of light scattering in plane-parallel media. In contrast, in the case of anisotropic scattering, the functions $Q_{mn}(\tau, t)$ depend not only on $|\tau - t|$, but also directly from the values of τ and t. Therefore, the system of integral Eqs. (75) is not reduced to the corresponding system of equations for plane-parallel media. Therefore, it is impossible to reduce the problem of radiative transfer in a sphere to the appropriate problem for a planar medium.

We note that the quantities $S_n(\tau;\tau_0)$ are determined by the well-known (see, for example, Bell and Glasstone 1970) recursion relation

$$n\tau^{n-1}\frac{\mathrm{d}}{\mathrm{d}\tau}\left[\frac{S_{n-1}(\tau;\tau_0)}{\tau^{n-1}}\right]+\frac{n+1}{\tau^{n+2}}\frac{\mathrm{d}}{\mathrm{d}\tau}\left[\tau^{n+2}S_{n+1}(\tau;\tau_0)\right]$$
$$=-(2n+1-\lambda x_n)S_n(\tau;\tau_0)\quad(n=0,1,\ldots). \tag{77}$$

Equations (73), (66), and (67) allow us to find the functions $S(\tau,\mu;\tau_0)$ and $I(\tau,\mu;\tau_0)$ if we know the quantities $S_n(\tau;\tau_0)$.

3.3 The Structure of Green's Function for Media with Spherical Symmetry

Following Kolesov (1985b), we will determine the structure of Green's function for an arbitrary spherically symmetric media.

We assume that the optical properties of these media are the same as the optical properties of an infinite homogeneous medium, as discussed in Sects. 1 and 2. The point positions will be characterized by optical distances τ from the center of the symmetry. As special cases of media with spherical symmetry, we consider a sphere of optical radius $\tau_0(0\le\tau\le\tau_0)$, a spherical shell with infinite optical thickness, i.e., an infinite medium $(\tau_*\le\tau<\infty)$ with a spherical cavity of optical radius τ_*, as well as a shell of finite optical thickness $\tau_0-\tau_*(\tau_*\le\tau\le\tau_0)$ bounded by spherical surfaces $\tau=\tau_*$ and $\tau=\tau_0$.

The radiation fields in such media with arbitrary spherically symmetric distributions of the sources are fully determined by the primary source function $D(\tau_1,\mu_1)$ and the Green's function $G(\tau,\mu;\tau_1,\mu_1)$.

As the Green's function $G_\infty(\tau,\mu;\tau_1,\mu_1)$ discussed in Sect. 2, the function $G(\tau,\mu;\tau_1,\mu_1)$ is also the solution of the homogeneous radiative transfer equation

$$\mu\frac{\partial G(\tau,\mu;\tau_1,\mu_1)}{\partial\tau}+\frac{1-\mu^2}{\tau}\cdot\frac{\partial G(\tau,\mu;\tau_1,\mu_1)}{\partial\mu}+G(\tau,\mu;\tau_1,\mu_1)$$
$$-\frac{\lambda}{2}\int_{-1}^{1}p(\mu,\mu')\,G(\tau,\mu';\tau_1,\mu_1)\mathrm{d}\mu'=0\quad(\tau\ne\tau_1). \tag{78}$$

This function satisfies the condition of the jump at the radiating surface

$$G(\tau_1+0,\mu;\tau_1,\mu_1)-G(\tau_1-0,\mu;\tau_1,\mu_1)=\frac{\delta(\mu-\mu_1)}{2\pi\tau_1^2\mu} \tag{79}$$

and the condition of reciprocity

$$G(\tau, \mu; \tau_1, \mu_1) = G(\tau_1, -\mu_1; \tau, -\mu). \tag{80}$$

These properties of the Green's function are common to all media of the spherical geometry and do not depend on the location of the boundary surfaces.

Jump condition (79) is valid for the Green's function $G_\infty(\tau, \mu; \tau_1, \mu_1)$ as well as for the function $G(\tau, \mu; \tau_1, \mu_1)$ in the cases of any media with spherical symmetry. Therefore, $G(\tau, \mu; \tau_1, \mu_1)$ can be represented as the sum $G_\infty(\tau, \mu; \tau_1, \mu_1)$ and some continuous on the radiating surface $\tau = \tau_1$ function. Since $G(\tau, \mu; \tau_1, \mu_1)$ and $G_\infty(\tau, \mu; \tau_1, \mu_1)$ are particular solutions of Eq. (78), then their difference $G(\tau, \mu; \tau_1, \mu_1) - G(\tau, \mu; \tau_1, \mu_1)$ is also a particular solution of this equation, i.e., the analytical expression of this difference must contain the eigenfunctions $f(\tau, \mu, \nu)$. From the condition of reciprocity (80), it follows that this expression must also contain the eigenfunctions $f(\tau_1, -\mu_1, \zeta) = f^*(\tau_1, \mu_1, -\zeta)$. Hence, this difference is a linear combination of the products of $f(\tau, \mu, \nu) f^*(\tau_1, \mu_1, -\zeta)$. Therefore, the structure of the Green's function for spherically symmetric media has the form

$$G(\tau, \mu; \tau_1, \mu_1) = G_\infty(\tau, \mu; \tau_1, \mu_1) - \frac{1}{2\pi} \mathbb{S}_\nu \mathbb{S}_\zeta c(\nu, \zeta) f(\tau, \mu, \nu) f^*(\tau_1, \mu_1, -\zeta). \tag{81}$$

In accordance with condition (80), the quantity $c(\nu, \zeta)$ is a symmetric function of ν and ζ, i.e., $c(\nu, \zeta) = c(\zeta, \nu)$.

Expression (81) gives a general form of the Green's functions for any spherically symmetric media. When you build this function for a particular medium, it is necessary to find the coefficients $c(\nu, \zeta)$, using the boundary conditions on the surfaces, in the center of symmetry, and at infinity. The expression for $c(\nu, \zeta)$ contains as parameters the optical distances of the boundaries of this media from the center of symmetry. For example, in the case of a sphere, quantities of $c(\nu, \zeta)$ depend on its optical radius τ_0, and in the case of a spherical shell they depend on the inner τ_* and external τ_0 optical radii. If we are interested in solving the radiative transfer equation, regular at the origin, we should put $c(\nu, \zeta) = 0$ if $\nu > 0$ or $\zeta > 0$. If we build solutions tending to zero at infinity, it is necessary to assume that $c(\nu, \zeta) = 0$ if $\nu < 0$ or $\zeta < 0$.

3.4 Integral Relations Between the Green's Functions

In the study of radiation fields in various media with spherical symmetry, it is advisable to use relations between the Green's functions for these media and for an infinite medium with a spherically symmetric distribution of the sources. The use of these relations allows using the results of the exact analytical theory developed for infinite media with spherical symmetry for studying the radiation fields in these media.

Before considering such relations for the media of spherical geometry, we give similar relations between the averaged over the azimuth Green's functions for a planar layer and for an infinite medium with planar symmetry. These relations are used for studying an asymptotic light regime in the outer layers of a sphere of large optical radius and of optically thick spherical shells.

The relations between the Green's functions for planar media were investigated by Domke and Ivanov (1975), Mnatsakanian (1976), Ivanov and Volkov (1979).

First of all, let us consider a planar homogeneous layer of optical thickness $\tau_0 - \tau_*$. The position of the points of the medium will be characterized by optical distances t from some arbitrary primary plane. The direction of propagation of radiation will be described by an angle $\arccos \eta$ between this direction and the normal to the primary plane. The boundary surface of the media will be considered parallel to this plane and to the planar radiation source.

Such planar layer can be represented as a part of an infinite medium bounded by the planes $t = \tau_0$ and $t = \tau_*(\tau_0 > \tau_*)$. We denote the averaged over the azimuth Green's function for this layer by $\bar{G}(t, \eta; t_1, \eta_1; \tau_0, \tau_*)$. Here $\tau_* \le t \le \tau_0$, $\tau_* \le t_1 \le \tau_0$, $-1 \le \eta \le 1$, $-1 \le \eta_1 \le 1$. When $\tau_* = -\infty, t \le \tau_0$ and when $\tau_0 = +\infty, t \ge \tau_*$, we have planar semi-infinite media. The Green's function $\bar{G}_\infty(t, \eta; t_1, \eta_1)$ for an infinite medium with planar symmetry was introduced in Sect. 1.

Let us write the relations between the functions $\bar{G}_\infty(t, \eta; t_1, \eta_1)$ and $\bar{G}(t, \eta; t_1, \eta_1; \tau_0, \tau_*)$. Thus, we shall proceed from the probabilistic meaning of the Green's function. This function represents the probability density of the following events. Let the quantum of light propagate on the optical distance t_1 from the primary plane at an angle $\arccos \eta_1$ to the normal; then as a result of diffusion, it will be on an optical distance t from this plane and will be propagate at an angle $\arccos \eta$ to the normal. Let the quantum of light radiate on the optical distance t_1 from the primary plane and propagate in an infinite medium. It may reach to some point on optical distance t from this plane propagating in various directions. First of all, it may be in the process of diffusion only within the layer $\tau_* \le t \le \tau_0$(if the point of its radiation is inside the layer). Second, it may enter into this layer from the half-spaces $t > \tau_0$ or $t < \tau_*$ through the surfaces $t = \tau_0$ or $t = \tau_*$, respectively, moving in various directions. Taking into account this fact, we obtain the following integral relation between the Green's functions:

$$\begin{aligned}\bar{G}_\infty(t, \eta; t_1, \eta_1) &= \bar{G}(t, \eta; t_1, \eta_1)\Theta(\tau_0 - t)\Theta(t - \tau_*) \\ &+ 2\pi \int_0^1 \bar{G}(\tau_0, -\eta'; t_1, \eta_1; \tau_0, \tau_*)\, \bar{G}_\infty(t, \eta; \tau_0, -\eta')\eta' d\eta' \\ &+ 2\pi \int_0^1 \bar{G}(\tau_*, \eta'; t_1, \eta_1; \tau_0, \tau_*)\, \bar{G}_\infty(t, \eta; \tau_*, \eta')\eta' d\eta' \quad (\tau_* \le t \le \tau_0),\end{aligned} \tag{82}$$

where $\Theta(x)$ is theta-function, i.e., $\Theta(x) = 1$ for $x > 0$, $\Theta(x) = 0$ for $x < 0$.

Let $\tau_* \le t_1 \le \tau_0$. A quantum of light, diffusing in an infinite medium, can leave the layer through the surfaces $t = \tau_*$ or $t = \tau_0$, moving at various angles to the normals, and then appear at a point located on an optical distance t from the primary plane. Hence, we have another integral relation between the Green's functions:

$$\begin{aligned}\bar{G}_\infty(t,\eta;t_1,\eta_1) &= \bar{G}(t,\eta;t_1,\eta_1;\tau_0,\tau_*)\Theta(\tau_0 - t)\Theta(t-\tau_*)\\ &+ 2\pi\int_0^1 \bar{G}(t,\eta;\tau_0,-\eta';\tau_0,\tau_*)\,\bar{G}_\infty(\tau_0,-\eta';t_1,\eta_1)\eta'\mathrm{d}\eta'\\ &+ 2\pi\int_0^1 \bar{G}(t,\eta;\tau_*,\eta';\tau_0,\tau_*)\,\bar{G}_\infty(\tau_*,\eta';t_1,\eta_1)\eta'\mathrm{d}\eta' \quad (\tau_* \le t_1 \le \tau_0).\end{aligned} \tag{83}$$

Taking into account the condition of reciprocity (80) for the function $\bar{G}_\infty(t,\eta;t_1,\eta_1)$ and a similar condition for $\bar{G}(t,\eta;t_1,\eta_1;\tau_0,\tau_*)$, we see that Eqs. (82) and (83) are equivalent, if inequalities $\tau_* \le \mathrm{t} \le \tau_0$ and $\tau_* \le t_1 \le \tau_0$ are performed simultaneously.

Particular cases of Eqs. (82) and (83) are the same relations between the Green's functions for a semi-infinite medium and for an infinite medium. These relations are obtained if the second or the third terms in the right parts of Eqs. (82) and (83) are equal to zero in the cases of $\tau_0 = +\infty$ or $\tau_* = -\infty$, respectively.

Relations (82), (83), and more general invariance relations were also obtained; for example, in works of Pikidjian (1978), Rogovtsov (1980), Rogovtsov and Samson (1985).

Consider now the integral relation between the Green's functions in the case of a sphere.

We denote by $G_i(\tau,\mu;\tau_1,\mu_1)$ and $G_e(\tau,\mu;\tau_1,\mu_1)$ the Green's functions for "internal" and "external" problems of the theory of radiation transfer in a sphere, respectively. The first of these functions describes the light regime in the sphere of optical radius τ_0. The second of these functions attribute to the case of an infinite medium with a spherically absolutely black cavity of optical radius τ_0. We assume that the external radiation illuminating the boundary surface of these media is missing. Then the boundary conditions are written in the form

$$G_i(\tau_0,\mu;\tau_1,\mu_1) = 0 \quad (-1 \le \mu \le 0), \tag{84}$$

$$G_e(\tau_0,\mu;\tau_1,\mu_1) = 0 \quad (0 < \mu \le 1). \tag{85}$$

Let imagine a spherical section in an infinite medium at a distance τ_0 from the center of symmetry. This section divides the infinite medium into two media. The

radiation fields in these media are described by the Green's functions $G_i(\tau, \mu; \tau_1, \mu_1)$ and $G_e(\tau, \mu; \tau_1, \mu_1)$. Taking into account the probabilistic meaning of the Green's functions and analyzing the possible trajectories of a quantum of light in the infinite medium, we find the integral relations between the Green's functions $G_i(\tau, \mu; \tau_1, \mu_1)$, $G_e(\tau, \mu; \tau_1, \mu_1)$ and $G_\infty(\tau, \mu; \tau_1, \mu_1)$. These relations have the form:

$$G_\infty(\tau, \mu; \tau_1, \mu_1) = G_i(\tau, \mu; \tau_1, \mu_1)\Theta(\tau_0 - \tau) + 2\pi\tau_0^2 \int_0^1 G_i(\tau_0, \mu'; \tau_1, \mu_1) G_\infty(\tau, \mu; \tau_0, \mu')\mu' \mathrm{d}\mu' \quad (0 \le \tau_1 \le \tau_0), \tag{86}$$

$$G_\infty(\tau, \mu; \tau_1, \mu_1) = G_e(\tau, \mu; \tau_1, \mu_1)\Theta(\tau - \tau_0) + 2\pi\tau_0^2 \int_0^1 G_e(\tau_0, -\mu'; \tau_1, \mu_1) G_\infty(\tau, \mu; \tau_0, -\mu')\mu' \, \mathrm{d}\mu' \quad (\tau_1 \ge \tau_0), \tag{87}$$

$$G_\infty(\tau, \mu; \tau_1, \mu_1) = G_i(\tau, \mu; \tau_1, \mu_1)\Theta(\tau_0 - \tau_1) + 2\pi\tau_0^2 \int_0^1 G_\infty(\tau_0, -\mu'; \tau_1, \mu_1) G_i(\tau, \mu; \tau_0, -\mu')\mu' \, \mathrm{d}\mu' \quad (0 \le \tau \le \tau_0), \tag{88}$$

$$G_\infty(\tau, \mu; \tau_1, \mu_1) = G_e(\tau, \mu; \tau_1, \mu_1)\Theta(\tau_1 - \tau_0) + 2\pi\tau_0^2 \int_0^1 G_\infty(\tau_0, \mu'; \tau_1, \mu_1) G_e(\tau, \mu; \tau_0, \mu')\mu' \, \mathrm{d}\mu' \quad (\tau \ge \tau_0). \tag{89}$$

Taking into account the condition of reciprocity for the Green's functions, we see that formulae (87) and (89) are equivalent if inequalities $\tau \ge \tau_0$ and $\tau_1 \ge \tau_0$ are performed simultaneously, and the formulae (86) and (88) are also equivalent if inequalities $0 \le \tau \le \tau_0$ and $0 \le \tau_1 \le \tau_0$ are also performed simultaneously.

In the same way, we may obtain the integral relation between the Green's functions for spherical shells and the infinite medium. These relations are similar to relations (82) and (83) for a planar layer.

3.5 Eigenfunctions Biorthogonal at the Half Interval of the Variation of the Angular Variable

When analyzing the radiative transfer problems in planar media using the Case method, the coefficients of the expansions of the Green's functions are found by solving the linear singular equations. The theory of such equations is described, in particular, in Muskhelishvili (1962), Gakhov (1966), Gakhov and Cherskij (1978). These equations result from the use of the property of biorthogonality for the eigenfunctions with respect to the conjugate eigenfunctions with the weight $\eta H(\eta)$ in the half intervals of the variation of an angular variable η $(0 \le \eta \le 1$ or $-1 \le \eta \le 0)$. Here $H(\eta)$ is H-function, introduced in the theory by Chandrasekhar (1960).

As it was noted in the literature (see, for example, Laletin 1974b), the drawback of this method is that it cannot be extended to the cases of more complicated geometries, in particular, to the practically important cases of media with spherical or cylindrical symmetry. The reason for this is that the type of singularity of eigenfunctions for nonplanar media other than for planar ones. The validity of the invariance principle for planar media leads to eigenfunctions containing the delta-function. The presence of this delta-function allows to obtain the linear singular integral equations. In the cases of spherical and cylindrical geometries, the eigenfunctions have a singularity of a different type. They are expressed in the form of divergent series in the spherical functions. As a result, it is not possible to reduce the problem of the radiation transfer in these media to the solution of the linear singular integral equations.

Domke (1983) formulated in another form the condition of biorthogonality for eigenfunctions $i(\eta, \nu)$ on the half of the interval $[-1,1]$. He showed that these functions are biorthogonal on the interval $0 \le \eta \le 1$ with weight η with respect to the eigenfunctions $u(\eta, \nu)$ of a generalized Milne problem (see, Case and Zweifel 1967), i.e.,

$$2\int_0^1 u(\eta, \nu) i(\eta, \zeta) \eta \, \mathrm{d}\eta = \delta(\nu, \zeta). \tag{90}$$

This condition allows the generalization for the case of media of nonplanar geometry.

In Kolesov (1985b) it were derived the eigenfunctions $U(\tau_0, \mu, \nu)$, which are an analogue of the function $u(\eta, \nu)$, and conjugate function $U(\tau_0, \mu, \nu) = U(\tau_0, -\mu, -\nu)$. The surface Green's functions $G_i(\tau, \mu; \tau_0, -\mu_1)$ and $G_e(\tau, \mu; \tau_0, \mu_1)$ are expressed through these eigenfunctions:

$$G_i(\tau, \mu; \tau_0, -\mu_1) = -\frac{1}{2\pi} \mathbb{S}_{\nu > 0} \frac{\nu^2}{N(\nu)} U(\tau_0, \mu_1, \nu) f^*(\tau, -\mu, \nu), \tag{91}$$

$$G_e(\tau,\mu;\tau_0,\mu_1) = -\frac{1}{2\pi}\mathbb{S}_{\nu>0}\frac{\nu^2}{N(\nu)}U^*(\tau_0,\mu_1,\nu)f(\tau,\mu,\nu). \tag{92}$$

The functions $U(\tau_0,\mu,\nu)$ and $U^*(\tau_0,\mu,\nu)$ are associated with the Case eigenfunctions $f(\tau_0,\mu,\nu)$ and the coefficients $\rho_i(\mu,\mu_1;\tau_0)$ and $\rho_e(\mu,\mu_1;\tau_0)$ of the reflection of light from a sphere and an infinite medium with a cavity:

$$\begin{aligned} U(\tau_0,\mu_1,\nu) = {} & f(\tau_0,\mu_1,\nu) - f(\tau_0,-\mu_1,\nu)\mathrm{e}^{-2\tau_0\mu_1} \\ & - 2\int_0^1 f(\tau_0,-\mu,\nu)\rho_i(\mu,\mu_1;\tau_0)\mu\,\mathrm{d}\mu, \end{aligned} \tag{93}$$

$$\begin{aligned} U^*(\tau_0,\mu_1,\nu) = {} & f^*(\tau_0,\mu_1,\nu) \\ & - 2\int_0^1 f^*(\tau_0,-\mu,\nu)\rho_e(\mu,\mu_1;\tau_0)\mu\,\mathrm{d}\mu, \end{aligned} \tag{94}$$

where

$$\rho_i(\mu,\mu_1;\tau_0) = \pi\tau_0^2 G_i(\tau_0,\mu;\tau_0,-\mu_1) - \frac{1}{2\mu}\delta(\mu-\mu_1)\mathrm{e}^{-2\tau_0\mu}, \tag{95}$$

$$\rho_e(\mu,\mu_1;\tau_0) = \pi\tau_0^2 G_e(\tau_0,-\mu;\tau_0,\mu_1). \tag{96}$$

Conditions of biorthogonality for the system of functions $U(\tau_0,\mu_1,\nu)$ and $U^*(\tau_0,\mu_1,\nu)$ can be written as

$$\int_0^1 U(\tau_0,\mu,\nu)f^*(\tau_0,\mu,\zeta)\mu\,\mathrm{d}\mu = -\frac{N(\nu)}{\tau_0^2\nu^2}\delta(\nu,\zeta), \tag{97}$$

$$\int_0^1 U^*(\tau_0,\mu,\nu)f(\tau_0,\mu,\zeta)\mu\,\mathrm{d}\mu = -\frac{N(\nu)}{\tau_0^2\nu^2}\delta(\nu,\zeta). \tag{98}$$

We note that if the functions $U(\tau_0,\mu_1,\nu)$ and $U^*(\tau_0,\mu_1,\nu)$ are used for solving the problems of radiative transfer in a media with spherical symmetry, you should preliminarily find the quantities $\rho_i(\mu,\mu_1;\tau_0)$ and $\rho_e(\mu,\mu_1;\tau_0)$.

3.6 Coefficients of Reflection and Transmission of Light

Coefficients of the diffuse reflection and transmission of light by various media are of special interest for the radiative transfer theory. The observed intensity of radiation is expressed through these quantities in many cases. For plane-parallel media, these quantities are investigated quite thoroughly (see, for example, Sobolev 1975). The coefficients of the reflection of light by a sphere and a spherical shell and coefficient of the light transmission by a spherical shell were studied by Bellman et al. (1959, 1960, 1966, 1968), Bellman and Kalaba (1965) and by Ueno et al. (1971) using the invariant embedding method. For these quantities, integro-differential equations were derived that contain derivatives with respect to the optical radius of a sphere or optical distances of the inner and outer boundaries of a spherical shell from the center of symmetry. However, analytical solutions of these equations have not been obtained.

The integral equations for these quantities were derived in Kolesov (1985b). These equations allow to find those quantities in condition of known light regime in an infinite medium with a spherical distribution of radiation sources.

Suggesting that in formula (86) $\tau = \tau_1 = \tau_0 - 0$, $0 \le \mu \le 1$, $-1 \le \mu \le 0$, and in formula (87) $\tau = \tau_1 = \tau_0 + 0$, $-1 \le \mu \le 0$, $0 \le \mu_1 \le 1$ and using relations (95), (96), and (101), we obtain the following linear integral equations for the reflection coefficients $\rho_i(\mu, \mu_1; \tau_0) = \rho_i(\mu_1, \mu; \tau_0)$ and $\rho_e(\mu, \mu_1; \tau_0) = \rho_e(\mu_1, \mu; \tau_0)$:

$$\rho_i(\mu, \mu_1; \tau_0) = \rho_\infty(\mu, \mu_1; \tau_0) - \rho_\infty(-\mu, \mu_1; \tau_0)\mathrm{e}^{-2\tau_0\mu} - 2\int_0^1 \rho_\infty(-\mu', \mu_1; \tau_0)\rho_i(\mu, \mu'; \tau_0)\mu'\,\mathrm{d}\mu', \tag{99}$$

$$\rho_e(\mu, \mu_1; \tau_0) = \rho_\infty(-\mu, -\mu_1; \tau_0) - 2\int_0^1 \rho_\infty(\mu', -\mu_1; \tau_0)\rho_e(\mu, \mu'; \tau_0)\mu'\,\mathrm{d}\mu'. \tag{100}$$

The reflection coefficient $\rho_\infty(\mu, \mu_1; \tau_0)$ for an infinite medium is associated with the corresponding Green's function by the relationships

$$\begin{aligned}\rho_\infty(\mu, \mu_1; \tau_0) &= \pi\tau_0^2 G_\infty(\tau_0 - 0, \mu; \tau_1, -\mu_1) \\ &\quad - \frac{\Theta(\mu_1)}{2|\mu|}\left[\delta(\mu - \mu_1)\mathrm{e}^{-2\tau_0\mu} + \delta(\mu + \mu_1)\right] \\ &= \pi\tau_0^2 G_\infty(\tau_0 + 0, \mu; \tau_0, -\mu_1) - \frac{\Theta(\mu_1)}{2|\mu|}\delta(\mu - \mu_1)\mathrm{e}^{-2\tau_0\mu}.\end{aligned} \tag{101}$$

Using the solution of the linear inhomogeneous equation describing the Green's function, we obtain the following expression for the quantity $\rho_\infty(\mu, \mu_1; \tau_0)$:

$$\rho_\infty(\mu, \mu_1; \tau_0) = -\frac{\lambda}{4}\tau_0^2 \sum_{n=0}^{N} x_n \mathbb{S}_{\nu>0} \frac{\nu^2}{N(\nu)} f(\tau_0, \mu_1, -\nu) \times \int_0^\infty f_n\left(\sqrt{\tau_0^2 + 2\tau_0\mu_1 s + s^2}, -\nu\right) P_n\left(\frac{\tau_0\mu + s}{\sqrt{\tau_0^2 + 2\tau_0\mu_1 s + s^2}}\right) \mathrm{e}^{-s}\mathrm{d}s. \tag{102}$$

Side by side with the diffuse reflection coefficient $\rho_\infty(\mu, \mu_1; \tau_0)$ the diffuse transmission coefficient $\sigma_\infty(\mu, \mu_1; \tau_0, \tau_*)$ is introduced, which is associated with the Green's function by the relationship

$$\sigma_\infty(\mu, \mu_1; \tau_0, \tau_*) = \pi\tau_0\tau_* G_\infty(\tau_0, \mu; \tau_*, \mu_1) - \frac{\tau_*}{2\tau_0\mu}\delta\left(\mu - \sqrt{1 - \frac{\tau_*^2}{\tau_0^2}(1 - \mu_1^2)}\right)\mathrm{e}^{-\tau_0\mu + \tau_1\mu_1}. \tag{103}$$

We find the following expression for this quantity:

$$\sigma_\infty(\mu, \mu_1; \tau_0, \tau_*) = -\frac{\lambda}{4}\tau_0\tau_* \sum_{n=0}^{N} x_n \mathbb{S}_{\nu>0} \frac{\nu^2}{N(\nu)} f^*(\tau_*, \mu_1, \nu) \times \int_0^\infty f_n\left(\sqrt{\tau_0^2 - 2\tau_0\mu s + s^2}, \nu\right) P_n\left(\frac{\tau_0\mu - s}{\sqrt{\tau_0^2 - 2\tau_0\mu s + s^2}}\right) \mathrm{e}^{-s}\,\mathrm{d}s. \tag{104}$$

Calculating the quantities $\rho_\infty(\mu, \mu_1; \tau_0)$, $\rho_\infty(-\mu, \mu_1; \tau_0)$, $\rho_\infty(\mu, -\mu_1; \tau_0)$, and $\rho_\infty(-\mu, -\mu_1; \tau_0)$ by expression (102) under conditions $0 \le \mu \le 1$ and $0 \le \mu_1 \le 1$, you can find the reflection coefficients $\rho_i(\mu, \mu_1; \tau_0)$ and $\rho_e(\mu, \mu_1; \tau_0)$ by numerical integration of Eqs. (99) and (100).

Ivanov (1976a, b) obtained an integral equation for the coefficient of the light reflection from a semi-infinite medium. This equation is similar to the Eqs. (99) and (100).

Integral equations describing the coefficients of the reflection and transmission of light by spherical shells of finite or infinite optical thickness with a completely black or a completely transparent spherical cavity were derived by Kolesov (1985c). These equations may be solved by numerical methods if the quantities $\rho_\infty(\mu, \mu_1; \tau_0)$ and $\sigma_\infty(\mu, \mu_1; \tau_*, \tau_0)$ are known.

A simple efficient numerical algorithm for computing bidirectional reflectance of a semi-infinite homogeneous medium is described by Mishchenko et al. (1999). This algorithm was generalized on the case of polarized radiation by Mishchenko et al. (2015).

4 Asymptotic Formulae of the Theory of Radiative Transfer in an Infinite Medium, a Sphere, and a Spherical Shell

In the radiation transfer theory it is comfortable to use simple asymptotic expressions for various characteristics of the radiation field to simplify calculations. Asymptotic expressions for these quantities are most commonly used in two practically important cases. First, it is the case of media of large optical sizes, when points of the media are located at large optical distances from sources of radiation. Second, it is the case of radiation fields with almost conservative scattering, when the albedo of single scattering is close to unity.

Let us at the beginning consider the first case (see Kolesov 1984b, 1985a).

4.1 The Asymptotic Light Regime in an Infinite Medium Far from a Point Source

In the derivation of asymptotic expressions for functions $S_\infty(\tau,\mu)$ and $I_\infty(\tau,\mu)$ when $\tau \gg 1$, we use the fact that the behavior of the functions $S_n^\infty(\tau)$ is defined by reintegrating term in the right part of expression (33), containing the largest root $\nu_1 = \frac{1}{k}$ of characteristic equation (12). You can neglect radiation received at a given point of the medium directly from the source. This results in the following asymptotic expression:

$$\begin{aligned} S_n^\infty(\tau) &= \frac{\lambda L\alpha^2}{16\pi^2}\cdot\frac{f_n\left(\tau,\frac{1}{k}\right)}{N\left(\frac{1}{k}\right)} \\ &\approx \frac{\lambda L\alpha^2}{16\pi^2}\cdot\frac{R_n\left(\frac{1}{k}\right)}{N\left(\frac{1}{k}\right)}\cdot\frac{k}{\tau}\sqrt{\frac{2k\tau}{\pi}}K_{n+\frac{1}{2}}(k\tau), \quad \tau \gg 1. \end{aligned} \tag{105}$$

When $k\tau \gg 1$, this formula is simplified and takes the form

$$S_n^\infty(\tau) \approx \frac{\lambda L\alpha^2}{16\pi^2}\cdot\frac{kR_n\left(\frac{1}{k}\right)}{N\left(\frac{1}{k}\right)}\cdot\frac{\mathrm{e}^{-k\tau}}{\tau}, \quad k\tau \gg 1. \tag{106}$$

Substitution (106) in (29) leads to the asymptotic expression for the source function:

$$S_\infty(\tau,\mu) \approx \frac{\lambda L\alpha^2}{16\pi^2}\cdot\frac{k}{N\left(\frac{1}{k}\right)}\cdot\sum_{n=0}^{N} x_n R_n\left(\frac{1}{k}\right)P_n(\mu)\cdot\frac{\mathrm{e}^{-k\tau}}{\tau}, \quad k\tau \gg 1. \tag{107}$$

In the case of isotropic scattering, this formula was obtained by Ambartsumyan (1945).

When $\tau \gg 1$, the simple asymptotic expression for $I_\infty(\tau, \mu)$ takes place (see Dolin 1981)

$$I_\infty(\tau, \mu) \approx \frac{L\alpha^2 k}{2\pi^2 \lambda M} \cdot i(\mu) \cdot \frac{\mathrm{e}^{-k\tau}}{\tau}. \tag{108}$$

Here, the function $i(\mu)$ associated with the quantity $R\left(\mu, \frac{1}{k}\right)$ by the relation

$$R\left(\mu, \frac{1}{k}\right) = \frac{\lambda}{2} \cdot i(\mu), \tag{109}$$

and the constant quantity M is related with the normalization integral $N\left(\frac{1}{k}\right)$ by the relationship

$$N\left(\frac{1}{k}\right) = \frac{\lambda^2}{8} \cdot M. \tag{110}$$

Substituting expressions (105) and (106) into Eq. (28), we obtain the following asymptotic expressions for the intensity of radiation:

$$I_\infty(\tau, \mu) \approx \frac{\lambda L\alpha^2}{16\pi^2 N\left(\frac{1}{k}\right)} \sum_{n=0}^{N} x_n \int_0^\infty f_n\left(\sqrt{\tau^2 - 2\tau\mu s + s^2}, \frac{1}{k}\right) P_n\left(\frac{\tau\mu - s}{\sqrt{\tau^2 - 2\tau\mu s + s^2}}\right) \mathrm{e}^{-s}\, \mathrm{d}s \quad (\tau \gg 1), \tag{111}$$

$$I_\infty(\tau, \mu) \approx \frac{\lambda L\alpha^2}{16\pi^2 N\left(\frac{1}{k}\right)} \times \sum_{n=0}^{N} x_n R_n\left(\frac{1}{k}\right) \int_0^\infty P_n\left(\frac{\tau\mu - s}{\sqrt{\tau^2 - 2\tau\mu s + s^2}}\right) \mathrm{e}^{-k\sqrt{\tau^2 - 2\tau\mu s + s^2} - s} \frac{\mathrm{d}s}{\sqrt{\tau^2 - 2\tau\mu s + s^2}} \quad (k\tau \gg 1). \tag{112}$$

4.2 *The Asymptotic Expression of Green's Function for an Infinite Medium with a Spherically Symmetric Distribution of Sources*

The asymptotic formulae connecting the Green's functions $G_\infty(\tau, \mu; \tau_1, \mu_1)$ and $\bar{G}_\infty(\tau, \mu; \tau_1, \mu_1)$ when $\tau \gg 1$ and $\tau_1 \gg 1$ were derived by Kolesov (1984b). The function $G_\infty(\tau, \mu; \tau_1, \mu_1)$ describes the light regime in an infinite homogeneous

medium with a spherical symmetric distribution of radiation sources. The Green's function $\bar{G}_\infty(\tau, \mu; \tau_1, \mu_1)$ corresponds to the case of an infinite medium illuminated by a planar source located in the plane $\tau = \tau_1$ and radiating in the direction forming an angle $\arccos \mu_1$ with the normal to the planar source. We note that for easy comparison of these Green's functions, optical distances τ in an infinite medium are counted from some arbitrary plane surface parallel to the planar source but not from this source. It is obvious that the values of the function $\bar{G}_\infty(\tau, \mu; \tau_1, \mu_1)$ depend on the difference $\tau - \tau_1$ but not separately from τ and from τ_1. The use of the same notations for function arguments of $\bar{G}_\infty(\tau, \mu; \tau_1, \mu_1)$ and $G_\infty(\tau, \mu; \tau_1, \mu_1)$ is justified by the fact that spherical layers have a little difference from corresponding planar layers when $\tau \gg 1$ and $\tau_1 \gg 1$.

Keeping in formulae (55) and (56) only the terms containing the discrete eigenvalue $\nu_1 = \frac{1}{k}$, substituting the expressions for the eigenfunctions into these equations and comparing the resulting formulae with expressions (17) and (18), we obtain the following asymptotic relation:

$$G_\infty(\tau, \mu; \tau_1, \mu_1) \approx \frac{\bar{G}_\infty(\tau, \mu; \tau_1, \mu_1)}{\tau\tau_1} - \frac{i(-\mu_1)i(\mu)}{\pi\tau\tau_1 M} \mathrm{e}^{-k(\tau+\tau_1)} \quad (\tau \gg 1, \tau_1 \gg 1). \tag{113}$$

The asymptotic expression for the Green's function $G_\infty(\tau, \mu; 0, 1)$ corresponding to the case of an isotropic central point source has the form:

$$G_\infty(\tau, \mu; 0, 1) \approx \frac{2k}{\pi\lambda M} i(\mu) \frac{\mathrm{e}^{-k\tau}}{\tau} (\tau \gg 1). \tag{114}$$

Kolesov (1990) studied the asymptotical light regime in a homogeneous infinite medium at large optical distances τ from an isotropic point source. An asymptotic expression for the radiation intensity was found in terms of inverse powers of τ. In the case of small true absorption, this quantity was expressed as a double power series of $^1/_\tau$ and the diffusion coefficient k.

4.3 *Asymptotic Radiation Fields in Outer Layers of a Sphere of Large Optical Radius Far from Sources*

Let us examine some works that have studied the asymptotic light regime in media with spherical symmetry.

Solving the problem of a point source in a homogeneous sphere with isotopic monochromatic scattering, Smith (1965) presented the source function in the form of infinite series that converge for sufficiently large values of the optical radius of the sphere.

Sobolev (1965) and Nagirner (1965) found asymptotic expressions for the source functions for outer layers of sphere of large optical radius and optically thick spherical shell. Germogenova (1966) investigated the asymptotic behavior of the radiation intensity in aspherical anisotropically scattering envelope surrounding a point source. Sobolev (1983) obtained asymptotic expressions for the radiation intensities emerging from a sphere or a spherical shell as well as for the light reflection coefficient of a sphere in the case of anisotropic scattering.

Ivanov (1969) and Nagirner (1972) studied the scattering of the resonance radiation in a homogeneous sphere of large optical radius.

Kolesov (1984b, 1985a) studied by the Case method the asymptotic regime in the outer layers of a sphere of large optical radius and of a spherical shell of large optical thickness in the case of anisotropic monochromatic scattering. In particular, he found asymptotic expressions for some characteristics of the radiation field.

Studying the asymptotic behavior of the Green's function $G_i(\tau, \mu; \tau_1, \mu_1)$ for a sphere of optical radius $\tau_0 \gg 1$, it is convenient to use integral relations (86) and (88) between functions $G_i(\tau, \mu; \tau_1, \mu_1)$ and $G_\infty(\tau, \mu; \tau_1, \mu_1)$, and the asymptotic expressions (115) and (116) for $G_\infty(\tau, \mu; \tau_1, \mu_1)$. In the case of a central point source ($\tau_1 = 0, \mu_1 = 1$), we have the following asymptotic expressions for the Green's function:

$$G_i(\tau, \mu; 0, 1) \approx \frac{2k}{\pi\lambda} \cdot \frac{I_M(\tau_0 - \tau, \mu)}{1 - N\mathrm{e}^{-2k\tau_0}} \cdot \frac{\mathrm{e}^{-k\tau_0}}{\tau} (\tau \gg 1,\ \tau_0 \gg 1,\ \lambda < 1), \tag{115}$$

$$G_i(\tau, \mu; 0, 1) \approx \frac{I_M(\tau_0 - \tau, \mu)}{\pi\tau\tau_0} (\tau \gg 1,\ \tau_0 \gg 1,\ \lambda = 1). \tag{116}$$

In these expressions, $I_M(\tau_0 - \tau_1\mu)$ is the relative intensity of radiation propagating at an angle $\arccos\mu$ to the external normal to the boundary surface on an optical depth $\tau_0 - \tau$ in a semi-infinite planar medium in the Milne problem(see Milne 1921), i.e., when radiation sources are sufficiently far away from the boundary surface of this medium. This function is given by the expression

$$I_M(\tau_0 - \tau, \mu) = \frac{i(\mu)}{M}\mathrm{e}^{k(\tau_0 - \tau)} - 2\pi \int_0^1 u(\mu')\bar{G}_\infty(\tau, \mu'; \tau_0, \mu)\mu' \,\mathrm{d}\mu', \tag{117}$$

where $u(\mu)$ is the relative intensity of the radiation emerging from the semi-infinite medium in the Milne problem. In particular,

$$I_M(0, \mu) = u(\mu). \tag{118}$$

In formula (115)

$$N = 2\int_0^1 u(\mu) i(-\mu)\mu \, \mathrm{d}\mu. \tag{119}$$

Let the radiation sources be located on the optical distance τ_1 from the center of a sphere and radiate in the direction forming an angle arccos μ_1 with the radius-vector. In this case, the expression for the Green functions has the following form:

$$G_i(\tau, \mu; \tau_1, \mu_1) \approx \frac{\bar{G}(\tau, \mu; \tau_1, \mu_1)}{\tau\tau_1} - \frac{M}{\pi\tau\tau_1} \cdot \frac{I_M(\tau_0 - \tau_1, -\mu_1) I_M(\tau_0 - \tau, \mu)}{1 - N\mathrm{e}^{-2k\tau_0}} \mathrm{e}^{-2k\tau_0} \quad (\tau \gg 1, \tau_1 \gg 1, \lambda < 1). \tag{120}$$

$$G_i(\tau, \mu; \tau_1, \mu_1) \approx \frac{\bar{G}(\tau, \mu; \tau_1, \mu_1)}{\tau\tau_1} \quad (\tau \gg 1,\ \tau_1 \gg 1,\ \lambda = 1). \tag{121}$$

Expressions (120) and (121) describe the asymptotic behavior of the Green's function of the so-called "internal" problems of the radiative transfer theory for a sphere.

Setting $\tau_1 = 0, \mu_1 = 1$, we obtain the following asymptotic expressions for the radiation intensity $I_i(\tau, \mu)$ in the surface layers of a sphere of optical radius $\tau_0 \gg 1$ with a central point source:

$$I_i(\tau, \mu) \approx \frac{L\alpha^2 k}{2\pi^2 \lambda} \cdot \frac{I_M(\tau_0 - \tau_1, \mu) I_M(\tau_0 - \tau, \mu)}{1 - N\mathrm{e}^{-2k\tau_0}} \cdot \frac{\mathrm{e}^{-k\tau_0}}{\tau} \quad (\tau \gg 1,\ \tau_1 \gg 1,\ \lambda < 1), \tag{122}$$

$$I_i(\tau, \mu) \approx \frac{L\alpha^2}{4\pi^2} \cdot \frac{I_M(\tau_0 - \tau, \mu)}{\tau\tau_0} \quad (\tau \gg 1,\ \tau_1 \gg 1,\ \lambda = 1). \tag{123}$$

In particular, when $\tau = \tau_0$, the asymptotic expressions for the intensity of the radiation emerging from a sphere may be represented as

$$I_i(\tau_0, \mu) \approx \frac{L\alpha^2 k}{2\pi^2 \lambda} \cdot \frac{u(\mu)}{1 - N\mathrm{e}^{-2k\tau_0}} \cdot \frac{\mathrm{e}^{-k\tau_0}}{\tau_0} \quad (\tau_0 \gg 1,\ \lambda < 1), \tag{124}$$

$$I_i(\tau_0, \mu) \approx \frac{L\alpha^2}{4\pi^2 \tau_0^2} \cdot u(\mu) \quad (\tau_0 \gg 1,\ \lambda = 1). \tag{125}$$

When sources of radiation are distributed in a sphere evenly with density D_0, the radiation intensity in the outer layers of this sphere is given by the asymptotic expression

$$I(\tau,\mu) \approx I_U(\tau_0-\tau,\mu) - \left[\frac{4}{\lambda k^2 \tau}\left(\mathrm{e}^{k\tau}-\mathrm{e}^{-k\tau}\right) - \frac{\alpha M}{1-\lambda}\mathrm{e}^{-k\tau_0}\right]\frac{D_0 I_M(\tau_0-\tau,\mu)}{1-N\mathrm{e}^{-2k\tau_0}}\mathrm{e}^{-k\tau_0}, \tag{126}$$

where $I_U(\tau,\mu)$ is the radiation intensity in a semi-infinite plane-parallel medium with uniformly distributed radiation sources. Here

$$\alpha = 2\int_0^1 u(\mu)\mu\,\mathrm{d}\mu. \tag{127}$$

The coefficient $\rho_i(\mu,\mu_1;\tau_0)$ of the light reflection from a sphere of large optical radius τ_0 satisfies the following asymptotic expressions:

$$\rho_i(\mu,\mu_1;\tau_0) \approx \bar{\rho}(\mu,\mu_1) - \frac{\mathrm{e}^{-k\tau_0}}{1-N\mathrm{e}^{-2k\tau_0}}u(\mu)u(\mu_1) \quad (\tau_0 \gg 1,\ \lambda<1), \tag{128}$$

$$\rho_i(\mu,\mu_1;\tau_0) \approx \bar{\rho}(\mu,\mu_1) \quad (\tau_0 \gg 1,\ \lambda=1). \tag{129}$$

Here $\bar{\rho}(\mu,\mu_1)$ is the coefficient of the light reflection from a semi-infinite planar medium. This quantity is investigated in detail, for example, in the books (Minin 1988; Sobolev 1975).

We also show the asymptotics of the radiation field in an infinite homogeneous medium with a spherical cavity bounded by absolutely black surface $\tau=\tau_0$. Such asymptotics corresponds to the Green's function $G_e(\tau,\mu;\tau_1,\mu_1)$ for the "external" problems of the radiative transfer theory in a sphere. The corresponding asymptotic expression is:

$$G_e(\tau,\mu;\tau_1,\mu_1) \approx \frac{\bar{G}(\tau,\mu;\tau_1,\mu_1)}{\tau\tau_1}, \quad \tau_0 \gg 1,\ \tau \geq \tau_0,\ \tau_1 \geq \tau_0,\ 0 \leq \lambda \leq 1, \tag{130}$$

where $\bar{G}(\tau,\mu;\tau_1,\mu_1)$ is the Green's function for a planar semi-infinite medium $\tau \geq \tau_0$.

4.4 The Asymptotic Behavior of the Radiation Field in an Optically Thick Spherical Envelope

Sobolev (1984) proposed an effective method of finding asymptotic formulae for the intensity of radiation emerging from the plane-parallel layers of large optical thickness for varions distributions of radiation sources. This method is based on the application of integral relations for the radiation intensity. It uses the fact that in simple cases the structure of asymptotic expressions are sometimes known from physical considerations. The problem is to determine factors included in these expressions. These factors are found by means of integral relations.

This method was applied in Kolesov (1984b, 1985a) to spheres with spherically symmetric source distributions. Let us consider this method.

The radiation intensity in a sphere or a spherical shell is determined by the integral-differential equation

$$\mu\frac{\partial I(\tau,\mu)}{\partial\tau}+\frac{1-\mu^2}{\tau}\frac{\partial I(\tau,\mu)}{\partial\mu}+I(\tau,\mu)$$
$$-\frac{\lambda}{2}\int_{-1}^{1}p(\mu,\mu')I(\tau,\mu')\mathrm{d}\mu'=D(\tau,\mu), \tag{131}$$

where $D(\tau,\mu)$ is the primary source function that describes the distribution of radiation sources in the medium. In Eq. (130), the quantity $I(\tau,\mu)$ is the full intensity of radiation, i.e., the sum of the intensities of radiation entered at a given point of the medium directly from a source and of diffuse radiation. However, as it was noted earlier, the full radiation intensity far from the radiation sources can be considered equal to the intensity of the diffuse radiation.

Let us consider the cases of a sphere of optical radius τ_0 and a shell bounded by spherical surfaces $\tau=\tau_*$ and $\tau=\tau_0$, where $\tau_*<\tau_0$. The central cavity will be considered fully transparent. As one of the boundary conditions imposed on the function $I(\tau,\mu)$, we may take the condition of the absence of external radiation falling on the surface $\tau=\tau_0$, i.e.,

$$I(\tau_0,\mu)=0 \quad \text{for } -1\le\mu\le 0. \tag{132}$$

In the case of a spherical shell as the second boundary condition, we have the condition

$$I(\tau_*,\mu)=I(\tau_*,-\mu) \quad \text{for } -1\le\mu<1. \tag{133}$$

This equality describes the transparency of the inner boundary surface of the shell and the absence of radiation sources within it. We note that the presence of external sources (for example, the central point source) can be taken into account by a suitable choice of the function $D(\tau,\mu)$.

To obtain the desired integral relations, we use the conjugate eigenfunctions $f_+(\tau,\mu)=f^*\left(\tau,\mu;\frac{1}{k}\right)$ and $f_-(\tau,\mu)=f^*\left(\tau,\mu;-\frac{1}{k}\right)$ of the homogeneous radiative transfer equation in spherically symmetric media.

After multiplying both parts of Eq. (131) by $f_\pm(\tau,\mu)$, and both parts of equation for functions $f_\pm(\tau,\mu)$ by $I(\tau,\mu)$, integrating the produced expressions over μ between -1 and 1 and also summarizing them, we obtain the equation

$$\frac{\mathrm{d}}{\mathrm{d}\tau}\left[\tau^2\int\limits_{-1}^{1} I(\tau,\mu)f_\pm(\tau,\mu)\mu\,\mathrm{d}\mu\right]=\tau^2\int\limits_{-1}^{1} D(\tau,\mu)f_\pm(\tau,\mu)\mathrm{d}\mu. \tag{134}$$

This equation is valid for any spherically symmetric media with an arbitrary spherically symmetric source distribution.

We note that the integrals in the left part of Eq. (134) are proportional to the coefficients $c\left(\tau,\pm\frac{1}{k}\right)$ of the radiation intensity expansion in the eigenfunctions. These coefficients correspond to the discrete eigenvalues $\pm\nu_1=\pm\frac{1}{k}$.

Taking into account the orthogonality of the eigenfunctions, we obtain

$$\tau^2\int\limits_{-1}^{1} I(\tau,\mu)f_\pm(\tau,\mu)\mu\,\mathrm{d}\mu=k^2N\left(\pm\frac{1}{k}\right)c\left(\tau,\pm\frac{1}{k}\right). \tag{135}$$

In the case of a sphere, we restrict ourselves using only the eigenfunction $f_+(\tau,\mu)$. Integrating Eq. (134), we use the condition that the left part of equality (135) tends to zero when $\tau\to 0$. Then we arrive at the integral relation

$$\tau^2\int\limits_{-1}^{1} I(\tau,\mu)f_+(\tau,\mu)\mu\,\mathrm{d}\mu=\int\limits_{0}^{\tau} t^2\,\mathrm{d}t\int\limits_{-1}^{1} D(t,\mu)f_+(t,\mu)\mathrm{d}\mu. \tag{136}$$

In the case of a spherical shell, the integration of Eq. (134) with boundary condition (133) gives the following integral relation:

$$\tau^2\int\limits_{-1}^{1} I(\tau,\mu)f_\pm(\tau,\mu)\mu\,\mathrm{d}\mu-\tau_*^2\int\limits_{0}^{1} I(\tau,\mu)[f_\pm(\tau_*,\mu)-f_\pm(\tau_*,-\mu)]\mu\,\mathrm{d}\mu$$
$$=\int\limits_{\tau_*}^{\tau} t^2\mathrm{d}t\int\limits_{-1}^{1} D(t,\mu)f_\pm(t,\mu)\mathrm{d}\mu. \tag{137}$$

Relations (136) and (137) are valid also when $\tau=\tau_0$. In this case, it is necessary to take into account the boundary condition (132).

Using the asymptotic expressions for the eigenfunctions, we obtain the following asymptotic integral relations:

$$\int_{-1}^{1} I(\tau,\mu)i(\mu)\mu\,\mathrm{d}\mu - \mathrm{e}^{-2k\tau}\int_{-1}^{1} I(\tau,\mu)i(-\mu)\mu\,\mathrm{d}\mu = -\frac{2}{\lambda k}\frac{\mathrm{e}^{-k\tau}}{\tau}\int_{0}^{\tau} t^2\,\mathrm{d}t\int_{-1}^{1} D(t,\mu)f_+(t,\mu)\mathrm{d}\mu \quad (\tau \gg 1) \tag{138}$$

in the case of a sphere of large radius, and

$$\tau\mathrm{e}^{k(\tau-\tau_*)}\int_{-1}^{1} I(\tau,\mu)i(\mu)\mu\,\mathrm{d}\mu + \tau\mathrm{e}^{-k(\tau-\tau_*)}\int_{-1}^{1} I(\tau,\mu)i(-\mu)\mu\,\mathrm{d}\mu = \int_{\tau_*}^{\tau} t^2\,\mathrm{d}t\int_{-1}^{1} D(t,\mu)\left[i(\mu)\mathrm{e}^{k(\tau-\tau_*)} + i(-\mu)\mathrm{e}^{-k(\tau-\tau_*)}\right]\mathrm{d}\mu \quad (\tau \gg 1, \tau_* \gg 1). \tag{139}$$

in the case of an optically thick spherical shell.

For a sphere by means of Eq. (138), we get the same asymptotic relations which have been obtained in the previous section of present article by another method.

Let consider the case of a spherical shell. If an isotropic point source of luminosity L is located in the center of the shell, then we have

$$D(\tau,\mu) = \frac{L\alpha^2}{8\pi^2\tau^2}\delta(\tau-\tau_*)\delta(\mu-1) \tag{140}$$

Substituting this function into integral relation (138), we obtain the asymptotic expression for the radiation intensity in outer layers of a spherical shell:

$$I(\tau,\mu) \approx \frac{L\alpha^2}{4\pi^2}[i(1)+i(-1)]\cdot\frac{I_M(\tau_0-\tau,\mu)}{1+Ne^{-2k\tau_0}}\cdot\frac{\mathrm{e}^{-k\tau_0}}{\tau_0\tau_*} \quad (\tau \gg \tau_0 \gg 1). \tag{141}$$

If the light sources are uniformly distributed in a spherical shell, then $D(\tau,\mu) = D_0 = \text{const}$. In this case we have

$$I(\tau,\mu) \approx I_U(\tau_0-\tau,\mu) - \left[\frac{4}{\lambda k^2\tau}\left(\mathrm{e}^{k\tau}-\mathrm{e}^{-k\tau}-2\right) + \frac{\alpha M}{1-\lambda}\mathrm{e}^{-k\tau_0}\right]\frac{D_0 I_M(\tau_0-\tau,\mu)}{1+N\mathrm{e}^{-2k\tau_0}}\mathrm{e}^{-k\tau_0} \quad (\tau_0 \gg \tau \gg 1), \tag{142}$$

where $I_U(\tau_0 - \tau, \mu)$ is the radiation intensity in a plane-parallel medium with uniformly distributed sources.

Assuming that the external surface $\tau = \tau_0$ of a spherical shell is illuminated by the conical source of luminosity L uniformly distributed over this surface with the unit density and emitting at an angle $\pi - \arccos \mu_0 (0 \le \mu_0 \le 1)$ to radius vector, we obtain the asymptotic expression for the coefficient of reflection of light from a spherical shell

$$\rho_s(\mu, \mu_0) \approx \bar{\rho}(\mu, \mu_0) - \frac{M e^{-k\tau_0}}{1 + N e^{-2k\tau_0}} u(\mu) u(\mu_0) \quad (\tau_0 \gg \tau_* \gg 1). \tag{143}$$

Developing asymptotic expressions in more complicated cases, one must take into account an influence of the whole spectrum of the characteristic equation on a solution of boundary value problems.

4.5 The Asymptotic Behavior of the Radiation Field Far Away from a Source in a Medium with a Low True Absorption

Simple asymptotic expressions for the characteristics of the radiation field may be obtained in the case of small true absorption of light ($1 - \lambda \ll 1$) in a media far away from sources of energy. We illustrate this with examples of an infinite homogeneous medium illuminated by a planar or a point source.

Such expressions may be derived using the known expressions of the quantities k and $N\left(\frac{1}{k}\right)$ in powers of $\sqrt{1 - \lambda}$ (see, for example, Minin 1988). With neglect of small terms of higher order than $\sqrt{1 - \lambda}$, we get

$$k \approx \sqrt{(3 - x_1)(1 - \lambda)} \quad (1 - \lambda \ll 1), \tag{144}$$

$$N\left(\frac{1}{k}\right) \approx \sqrt{\frac{1 - \lambda}{3 - x_1}} \quad (1 - \lambda \ll 1). \tag{145}$$

Let us find asymptotic expressions for the mean intensity and for the radiation flux using the expansion of the intensity in the divergent series in Legendre polynomials.

Let an infinite medium be illuminated by an isotropic planar source. The Green's function far away from the source (at $t \gg 1$) is given by expression (20) (for $t_1 = 0$), in which we take into account only the term corresponding the discrete eigenvalue $\nu_1 = \frac{1}{k}$, i.e.,

$$\bar{G}_\infty(t,\eta;0) \approx \frac{1}{4\pi}\cdot\frac{R\left(\eta,\frac{1}{k}\right)}{N\left(\frac{1}{k}\right)}\mathrm{e}^{-kt} \quad (t \gg 1). \tag{146}$$

The radiation intensity $\bar{I}_\infty(t,\eta)$ may be written in the form

$$\bar{I}_\infty(t,\eta) = lL\,\bar{G}_\infty\,(t,\eta;0). \tag{147}$$

Using expression (10) for the function $R\left(\eta,\frac{1}{k}\right)$, we obtain

$$\bar{I}_\infty(t,\eta) = \frac{lL}{8\pi}\sum_{n=0}^{\infty}(2n+1)\frac{R_n\left(\frac{1}{k}\right)}{N\left(\frac{1}{k}\right)}P_n(\eta)\mathrm{e}^{-kt}. \tag{148}$$

Then, using asymptotic expressions (144) and (145) for the quantities k and $N\left(\frac{1}{k}\right)$, we find the following expressions for the mean intensity $\bar{J}_\infty(t)$ and the radiation flux $\bar{H}_\infty(t)$:

$$\bar{J}_\infty(t) \approx \frac{1}{2}\int_{-1}^{1}\bar{I}_\infty(t,\eta)\,\mathrm{d}\eta = \frac{lL}{8\pi}\sqrt{\frac{1-\lambda}{3-x_1}}\mathrm{e}^{-\sqrt{(3-x_1)(1-\lambda)}t} \quad (t \gg 1, 1-\lambda \ll 1), \tag{149}$$

$$\bar{H}_\infty(t) \approx 2\pi\int_{-1}^{1}\bar{I}_\infty(t,\eta)\eta\mathrm{d}\eta = \frac{lL}{2}\mathrm{e}^{-\sqrt{(3-x_1)(1-\lambda)}t} \quad (t \gg 1, 1-\lambda \ll 1). \tag{150}$$

In the case of the central point source, assuming that $\tau \gg 1$, $1-\lambda \ll 1$ and using (37) and (38), we obtain the following asymptotic expressions for the functions $J_\infty(\tau)$ and $H_\infty(\tau)$:

$$J_\infty(\tau) \approx \frac{L\alpha^2}{16\pi^2\tau}(3-x_1)\mathrm{e}^{-\sqrt{(3-x_1)(1-\lambda)}\tau} \quad (\tau \gg 1, 1-\lambda \ll 1), \tag{151}$$

$$H_\infty(\tau) \approx \frac{L\alpha^2}{4\pi\tau^2}\left[1+\sqrt{(3-x_1)(1-\lambda)}\tau\right]\mathrm{e}^{-\sqrt{(3-x_1)(1-\lambda)}\tau} \quad (\tau \gg 1, 1-\lambda \ll 1). \tag{152}$$

It should be noted that the asymptotic expressions for the Green's function and the radiation intensity derived by the Case method does not describe the effect of peak intensity at $\mu = 1$. Therefore the errors of these expressions increase when values of μ approach to unity.

The real turbid media phase functions are sharply anisotropic and the corresponding characteristic equation can have numerous positive discrete roots. In this case to derive asymptotic expressions for radiation field parameters it is not

sufficient to use only the eigenfunction corresponding to the largest root $\nu_1 = \frac{1}{k}$. Konovalov (1974) pointed out that it is possible to obtain polyadic asymptotics for brightness coefficients taking into account contributions of some finite part of discrete roots. Rogovtsov (1986b, 2010), for example, suggested rigorous method of obtaining polyadic asymptotics.

5 Radiation Fields in Media with Cylindrical Symmetry

Most of the objects studied in astrophysics can be represented by models of plane-parallel or spherical media. However, some objects (sun spots, accretion disks and others) have an axial symmetry. If their length is considerably greater than the thickness, then these objects can be described by the model of an infinitely long cylinder. Such a model can be used, for example, for studying solar prominences, coronal rays, and so on. So the problem of light propagation in media with cylindrical symmetry is of particular astrophysical interest. A similar problem arising in the theory of neutron diffusion (see, for example, Laletin 1974a; Weinberg and Wigner 1958) is of interest for physics of nuclear reactors.

5.1 Basic Methods for Solving the Problems of the Radiative Transfer in Media with Cylindrical Symmetry

One-dimensional transfer problems with cylindrical symmetry is mathematically similar to those with spherical symmetry. Methods of solving spherical problems, a brief overview of which is given in Sect. 3 of the present article, are usually quite easily applied to cylindrical problems. We list the most frequently used methods.

Various approximate methods are often used for the solution of the radiative transfer and neutron diffusion problems in media with axial symmetry. The spherical harmonics method (P_N-approximation) has been described, for example, in the work by Davison (1957a), in books by Weinberg and Wigner (1958), Davison (1957b), Marchuk and Lebedev (1981), Davydov and Shikhov (1973). The Eddington method was used by Brand (1979), Heaslet and Warming (1966). The S_N-method, the double P_N-approximation, F_N-method and method of pseudo sources (G_N-approximation) have been applied by Shikhov and Troyanovski (1983), Drawbaugh and Noderer (1959), Siewert and Thomas (1985), and Laletin (1974b), respectively.

The integro-differential equations for the functions of light reflection and transmission for a homogeneous cylindrical shell were derived by the invariant imbedding method by Bellman and Kalaba (1957), Kalaba and Ruspini (1971), and Tsujita (1967, 1969).

Similar equations in an inhomogeneous cylindrical shell with anisotropic scattering were obtained by the probabilistic method in the article by Leong and Sen (1970). The non-stationary radiation field in the same shell with isotropic scattering was investigated using the same method by Leong (1972). In Apruzese (1981), Leong and Sen (1970, 1973b) this method was used for the study of the resonance radiative transfer with the frequency redistribution in media with cylindrical symmetry.

The integral equation of radiative transfer in a medium with cylindrical symmetry was solved by Leong and Sen (1971b) using the method of presenting its kernel as a sum of degenerate and small components, and by Kho and Sen (1972) using the expansion of the source function in a Neumann series.

Kuznetsov (1962) examined the radiative transfer in an inhomogeneous cylinder with a spherical phase function under the assumption of the nonmonochromatic radiation equilibrium. He got an integral equation describing the radial temperature distribution.

The numerical method, in particular, the MonteCarlo method is often used for calculation of the radiation fields in media of cylindrical geometry (see, for example, Avery et al. 1969; Peraiah 1973, 1975, 1978; Peraiah and Grant 1973).

A number of works devoted to the development of analytical methods which may be applied in the theory of radiative transfer problems in media with cylindrical symmetry. Hund (1968) obtained the exact solution of the problem of the line source in an infinite medium. Laletin (1966) found the cylindrically symmetric elementary solutions of the radiative transfer equation possessing cylindrical symmetry. He constructed the Green's function for an absorbing and isotropically scattering infinite medium with a cylindrically symmetric distribution of sources (see Laletin 1969a, 1974a).

Germogenova and Pavelyeva (1989), Pavelyeva (1990) considered the characteristics of radiation fields in extended cylindrical areas but without exact values of normalization factors. The values of these factors were founded by Rogovtsov (1990, 1992a).

Nagirner (1986, 1994a, b, 1995) investigated the process of the radiation transfer in spectral lines with isotropic scattering in the cases of complete radiation redistribution in frequency and without frequency change. He received asymptotic expressions for some characteristics of radiation fields. These expressions are valid at large optical distances from the symmetry axis.

Kolesov (1986) used the method of eigenfunctions to develop the radiative transfer theory in cylindrical absorbing and anisotropically scattering media. In the case of the infinite medium with cylindrically symmetric distribution of sources, the exact analytical expression for characteristics of the radiation field was derived. In the case of homogeneous media bounded by a cylindrical surface, the integral relations were derived. These relations connect the characteristics of radiation field in such media with the corresponding quantities for the infinite medium.

5.2 The Radiative Transfer Equation in Media with Cylindrical Symmetry

Let us consider an infinite medium with the same optical properties as properties of previously reviewed plane-parallel media and media with spherical symmetry. Let us assume that the distribution of sources depends on optical distances ρ from the axis of symmetry.

The direction of the radiation propagation at some point of the medium will be described by coordinates in a spherical system. The polar axis can be oriented either parallel or perpendicular to the symmetry axis. The polar and azimuthal angles we denote in the first case by $\arccos\xi$ and ψ while in the second case we denote them by $\arccos\mu$ and φ, respectively. The radiation intensity in these cases we denote by $\tilde{I}(\rho,\xi,\psi)$ and $\tilde{I}(\rho,\mu,\varphi)$.

The radiative transfer equation in the first case is written as

$$\sqrt{1-\xi^2}\left[\cos\psi\frac{\partial\tilde{I}(\rho,\xi,\psi)}{\partial\rho}-\frac{\sin\psi}{\rho}\cdot\frac{\partial\tilde{I}(\rho,\xi,\psi)}{\partial\psi}\right]+\tilde{I}(\rho,\xi,\psi)$$
$$-\frac{\lambda}{4\pi}\int_0^{2\pi}\mathrm{d}\psi'\int_{-1}^{1}x(\cos\gamma')\tilde{I}(\rho,\xi',\psi')\mathrm{d}\xi'=D(\rho,\xi,\psi),\qquad(153)$$

In the second case it has the form

$$\mu\frac{\partial\tilde{I}(\rho,\mu,\varphi)}{\partial\rho}+\frac{1-\mu^2}{\rho}\cos^2\varphi\cdot\frac{\partial\tilde{I}(\rho,\mu,\varphi)}{\partial\mu}+\frac{\mu}{2\rho}\sin 2\varphi\frac{\partial\tilde{I}(\rho,\mu,\varphi)}{\partial\varphi}$$
$$+\tilde{I}(\rho,\mu,\varphi)-\frac{\lambda}{4\pi}\int_0^{2\pi}\mathrm{d}\varphi'\int_{-1}^{1}x(\cos\gamma')\tilde{I}(\rho,\mu',\varphi')\mathrm{d}\mu'=D(\rho,\mu,\varphi).\qquad(154)$$

In these equations $D(\rho,\xi,\psi)$ or $D(\rho,\mu,\varphi)$ are the primary source functions. The cosine of the scattering angle γ' is given by the relations

$$\cos\gamma'=\xi\xi'+\sqrt{\left(1-\xi^2\right)\left(1-\xi'^2\right)}\cos(\psi-\psi')$$
$$=\eta\eta'+\sqrt{(1-\eta^2)(1-\eta'^2)}\cos(\varphi-\varphi').\qquad(155)$$

5.3 Eigenfunctions and Eigenvalues of the Homogeneous Radiative Transfer Equation in a Media with Cylindrical Symmetry

We denote solutions of homogeneous radiative transfer equations corresponding to Eqs. (153) and (154) by $F(\rho, \xi, \psi, \nu)$ and $f(\rho, \mu, \varphi, \nu)$, respectively. Here ν is the eigenvalue. Equations (153) and (154) are equivalent, since they are the same equation, but written in different coordinates. From their equivalence, it follows that

$$f(\rho, \mu, \varphi, \nu) = F\left(\rho, \sqrt{1-\mu^2}\sin\varphi, \operatorname{arctg}\frac{\sqrt{1-\mu^2}\cos\varphi}{\mu}, \nu\right), \tag{156}$$

$$F(\rho, \xi, \psi, \nu) = f\left(\rho, \sqrt{1-\xi^2}\cos\psi, \operatorname{arctg}\frac{\xi}{\sqrt{1-\xi^2}\sin\psi}, \nu\right). \tag{157}$$

The equation for $F(\rho, \xi, \psi, \nu)$ is simpler. Thus, using it, it is easier to find expressions for the eigenfunctions. The equation for $f(\rho, \mu, \varphi, \nu)$ asymptotically goes to the appropriate equation for the case of planar media when $\rho \gg 1$. Therefore, it is useful to find the asymptotics of the eigenfunctions.

The structure of the function $F(\rho, \xi, \psi, \nu)$ is the following:

$$\begin{aligned} &F(\rho, \xi, \psi, \nu) \\ &\quad = \frac{1}{4\pi}\sum_{m=0}^{\infty}(2-\delta_{m0})\cos m\psi \sum_{n=m}^{\infty}(2n+1)\frac{(n-m)!}{(n+m)!}F_n^m(\rho, \nu)P_n^m(\xi), \end{aligned} \tag{158}$$

where

$$F_n^m(\rho, \nu) = r_n^m(\nu)K_m\left(\frac{\rho}{\nu}\right) \tag{159}$$

or

$$F_n^m(\rho, \nu) = \pi r_n^m(\nu)I_m\left(-\frac{\rho}{\nu}\right). \tag{160}$$

Here $K_m(z)$ and $I_m(z)$ are the modified Bessel functions of the 1st kind and 3rd kind, with the integer index m.

The values of the Bessel functions $K_m(z)$ tend to zero as $z \to +\infty$. The functions $I_m(z)$ are regular at the origin. Consequently, the eigenfunctions $K_m\left(\frac{\rho}{\nu}\right)$ can be used when $\nu > 0$ to construct solutions decreasing as $\rho \to +\infty$, and the eigenfunctions $I_m\left(-\frac{\rho}{\nu}\right)$ are used when $\nu < 0$ to build solutions that are regular at $z = 0$.

The coefficients $r_n^m(\nu)$ may be found by means of the relations between them and the eigenfunctions $R(\mu, \varphi, \nu)$ for plane media:

$$r_n^m(\nu) = \int_0^{2\pi} \mathrm{d}\varphi \int_{-1}^{1} R(\mu, \varphi, \nu) P_n^m\left(\sqrt{1-\mu^2}\sin\varphi\right)\left(\cos\, m\,\mathrm{arctg}\,\frac{\sqrt{1-\mu^2}\cos\varphi}{\mu}\right)\mathrm{d}\mu, \tag{161}$$

where $P_n^m(z)$ are the associated Legendre functions.

The series (158) converges when $\nu<0$, e.g., when the coefficients $F_n^m(\rho,\nu)$ are determined by expression (160), and diverges when $\nu>0$, e.g., when these coefficients are given by expression (159). Thus, to compute the functions $F(\rho,\xi,\psi,\nu)$ when $\nu<0$, you may directly use expression (158), and if $\nu>0$, you should use the formal solution of the corresponding radiative transfer equation.

Along with the eigenfunctions $F(\rho,\xi,\psi,\nu)$, we have to use the conjugate functions $F^*(\rho,\xi,\psi,\nu)$ that

$$F^*(\rho,\xi,\psi,\nu) = F(\rho,-\xi,\pi+\psi,-\nu) \quad (0\le\psi\le\pi), \tag{162}$$

$$F^*(\rho,\xi,\psi,\nu) = F(\rho,-\xi,\pi-\psi,-\nu) \quad (\pi\le\psi\le 2\pi). \tag{163}$$

The eigenfunctions have the property of orthogonality:

$$\int_0^{2\pi}\cos\psi\,\mathrm{d}\psi\int_{-1}^{1}\sqrt{1-\xi^2}F(\rho,\xi,\psi,\nu)F^*(\rho,\xi,\psi,\zeta)\mathrm{d}\xi = \frac{\pi|\nu|}{2\rho}\bar{N}(\nu)\delta(\nu,\zeta), \tag{164}$$

$$\int_0^{2\pi}\mathrm{d}\varphi\int_{-1}^{1}\mu f(\rho,\mu,\varphi,\nu)f^*(\rho,\mu,\varphi,\zeta)\mathrm{d}\mu = \frac{\pi|\nu|}{2\rho}\bar{N}(\nu)\delta(\nu,\zeta), \tag{165}$$

where $f^*(\rho,\mu,\varphi,\xi) = f(\rho,-\mu,\varphi\pm\pi,-\xi)$.

We note that asymptotic expressions for functions $f(\rho,\mu,\varphi,\nu)$ and $F_n^m(\rho,\nu)$

$$f(\rho,\mu,\varphi,\nu) = \sqrt{\frac{\pi|\nu|}{2\rho}}R(\mu,\varphi,\nu)\mathrm{e}^{-\frac{\rho}{\nu}} \quad \left(\frac{\rho}{|\nu|}\gg 1\right), \tag{166}$$

$$F_n^m(\rho,\nu) = \sqrt{\frac{\pi|\nu|}{2\rho}}r_n^m(\nu)\mathrm{e}^{-\frac{\rho}{\nu}} \quad \left(\frac{\rho}{|\nu|}\gg 1\right). \tag{167}$$

are valid at large optical distances from the axis of symmetry.

Because systems of eigenfunctions $f(\rho,\mu,\varphi,\nu)$ and $F(\rho,\xi,\psi,\nu)$ have the properties of orthogonality and completeness, the radiative transfer problem solution with cylindrical symmetry may be found in the form of expansions on the system of functions $f(\rho,\mu,\varphi,\nu)$ or on the system of functions $F(\rho,\xi,\psi,\nu)$. In

particular, for the Green's function $G_\infty(\rho, \xi, \psi; \rho_1, \xi_1, \psi_1)$, we obtain the following expressions:

$$G_\infty(\rho, \xi, \psi; \rho_1, \xi_1, \psi_1) = \frac{2}{\pi}\mathbb{S}_{\nu>0}\frac{F^*(\rho_1, \xi_1, \psi_1, \nu)F(\rho, \xi, \psi, \nu)}{\nu\bar{N}(\nu)} \quad (\rho > \rho_1), \quad (168)$$

$$G_\infty(\rho, \xi, \psi; \rho_1, \xi_1, \psi_1) = \frac{2}{\pi}\mathbb{S}_{\nu>0}\frac{F^*(\rho_1, \xi_1, \psi_1, -\nu)F(\rho, \xi, \psi, -\nu)}{\nu\bar{N}(\nu)} \quad (\rho < \rho_1). \quad (169)$$

In the same way as in the case of an infinite medium with a planar or a point source of radiation (see Sect. 4.5), the expressions for the mean radiation intensity $\widetilde{J}_\infty(\tau)$ and for the radiation flux $\widetilde{H}_\infty(\tau)$ are obtained for the case of the small true absorption of light in an infinite homogeneous medium far away from an axial source:

$$\widetilde{J}_\infty(\tau) \approx \frac{lL}{16\pi^3}\sqrt{\frac{3-x_1}{1-\lambda}}K_0\left(\tau\sqrt{(3-x_1)(1-\lambda)}\right) \quad (k\tau \gg 1,\ 1-\lambda \ll 1), \quad (170)$$

$$\widetilde{H}_\infty(\tau) \approx \frac{lL}{4\pi^2}K_0\left((\tau\sqrt{(3-x_1)(1-\lambda)}\right) \quad (k\tau \gg 1,\ 1-\lambda \ll 1). \quad (171)$$

6 Nonstationary Radiation Fields

Fundamentals of the theory of the nonstationary radiative transfer are described in works by Sobolev (1952, 1963), Case and Zweifel (1967), and Minin (1988). Detailed overviews of the results obtained in this theory were published by Nagirner (1974) and Grinin (1994). The asymptotic expressions for the mean intensity and the radiation flux are obtained by Kolesov and Kropacheva (2013, 2014) in the case of the nonstationary radiation field in an infinite medium with a small true absorption. In these works it is considered the cases of planar, point, and axial radiation sources.

6.1 Basic Equations of the Theory of Nonstationary Radiative Transfer

Let us consider an infinite homogeneous medium characterized by the absorption coefficient α, the single scattering albedo λ and the phase function $x(\cos\gamma)$, where γ is the scattering angle. Nonstationary radiation fields are also characterized by the finite speed of light and definite durations of the light scattering process.

It is usually assumed that the probability of emission of a photon being in the absorbed state in the time interval from t to $t + \mathrm{d}t$ depends on t by the exponential law, i.e., it is proportional to $\mathrm{e}^{-\frac{t}{t_1}}\frac{\mathrm{d}t}{t_1}$. Here, t_1 is the mean time of the stay of a photon in the absorbed state.

The probability of the photon absorption while travelling after his radiation during an interval of time from t to $t + \mathrm{d}t$ depends on t also exponentially, i.e., it is proportional to $e^{-\frac{t}{t_2}}\frac{\mathrm{d}t}{t_2}$. Here $t_2 = \frac{1}{\alpha c}$ is the mean time of stay of a photon on the path between two consecutive scatterings, c is the speed of light.

In the simplest case, the quantities $\alpha, \lambda, x(\cos\gamma), t_1,$ and t_2 are considered to be independent of time and spatial coordinates.

Instead of the geometric distances r, the physical time t, and values t_1 and t_2, we use the corresponding dimensionless quantities

$$\tau = \alpha r, \quad u = \frac{t}{t_1 + t_2}, \quad \beta_1 = \frac{t_1}{t_1 + t_2}, \quad \beta_2 = \frac{t_2}{t_1 + t_2}. \tag{172}$$

If the medium has a plane-parallel geometry, then the equation of nonstationary radiative transfer is written in the form (see Sobolev 1963):

$$\eta \frac{\partial \bar{I}(\tau, \eta, u)}{\partial \tau} + \beta_2 \frac{\partial \bar{I}(\tau, \eta, u)}{\partial u} + \bar{I}(\tau, \eta, u) = \bar{B}(\tau, \eta, u), \tag{173}$$

where

$$\bar{B}(\tau, \eta, u) = \frac{\lambda}{2} \int_{-1}^{1} p(\eta, \eta')\mathrm{d}\eta' \int_{0}^{u} \bar{I}(\tau, \eta', u')\mathrm{e}^{-\frac{u-u'}{\beta_1}} \frac{\mathrm{d}u'}{\beta_1} + g(\tau, u). \tag{174}$$

In Eqs. (173) and (174), the function $\bar{I}(\tau, \eta, u)$ is the intensity of radiation at the dimensionless time u on the optical depth τ, the radiation is propagating in the direction which is forming an angle $\arccos\eta$ with the normal to the boundary surface. $\bar{B}(\tau, \eta, u)$ is the source function. $g(\eta, u)$ is the function describing the distribution of primary sources in the medium at time u and $p(\eta, \eta')$ is the phase function averaged over the azimuth.

In the case of a medium with spherical symmetry, this equation is written in the form:

$$\mu \frac{\partial I(\tau, \mu, u)}{\partial \tau} + \frac{1 - \mu^2}{\tau} \frac{\partial I(\tau, \mu, u)}{\partial \mu} + \beta_2 \frac{\partial I(\tau, \mu, u)}{\partial u} + I(\tau, \mu, u) = B(\tau, \mu, u), \tag{175}$$

where

$$B(\tau,\mu,u)=\frac{\lambda}{2}\int\limits_{-1}^{1}p(\mu,\mu')\mathrm{d}\mu'\int\limits_{0}^{u}I(\tau,\mu',u')\mathrm{e}^{-\dfrac{u-u'}{\beta_1}}\frac{\mathrm{d}u'}{\beta_1}+g(\tau,u). \tag{176}$$

Here τ is the optical distance from the center of symmetry and $\arccos\mu$ is the angle between the direction of light propagation and the radius vector.

When the medium has axial symmetry, the corresponding equation is written as:

$$\begin{aligned}\sqrt{1-\xi^2}\left[\cos\psi\frac{\partial\tilde{I}(\tau,\xi,\psi,u)}{\partial\tau}-\frac{\sin\psi}{\tau}\cdot\frac{\partial\tilde{I}(\tau,\xi,\psi,u)}{\partial\psi}\right]\\ +\beta_2\frac{\partial\bar{I}(\tau,\xi,\psi,u)}{\partial u}+\widetilde{I}(\tau,\xi,\psi,u)\\ =\frac{\lambda}{4\pi}\int\limits_{0}^{2\pi}\mathrm{d}\psi'\int\limits_{-1}^{1}x(\cos\gamma')\mathrm{d}\xi'\int\limits_{0}^{u}\widetilde{I}(\tau,\xi',\psi',u)\mathrm{e}^{-\dfrac{u-u'}{\beta_1}}\frac{\mathrm{d}u'}{\beta_1}+g(\tau,u),\end{aligned} \tag{177}$$

if the polar axis is directed parallel to the axis of symmetry. Here $\arccos\xi$ and ψ are polar and azimuthal angles. The quantity τ is the optical distance from the axis of symmetry.

If the polar axis is perpendicular to the symmetry axis, then the equation has the form:

$$\begin{aligned}\mu\frac{\partial\widetilde{I}(\tau,\mu,\varphi,u)}{\partial\tau}+\frac{1-\mu^2}{\tau}\cos^2\varphi\frac{\partial\widetilde{I}(\tau,\mu,\varphi,u)}{\partial\mu}\\ +\frac{\mu}{2\tau}\sin 2\varphi\frac{\partial\widetilde{I}(\tau,\mu,\varphi,u)}{\partial\varphi}+\beta_2\frac{\partial\widetilde{I}(\tau,\mu,\varphi,u)}{\partial u}\\ +\widetilde{I}(\tau,\mu,\varphi,u)=\frac{\lambda}{4\pi}\int\limits_{0}^{2\pi}\mathrm{d}\varphi'\int\limits_{-1}^{1}x(\cos\gamma')\mathrm{d}\mu'\int\limits_{0}^{u}\widetilde{I}(\tau,\mu',\varphi',u')\mathrm{e}^{-\dfrac{u-u'}{\beta_1}}\frac{\mathrm{d}u'}{\beta_1}+g(\tau,u),\end{aligned} \tag{178}$$

where $\arccos\mu$ and φ are appropriate polar and azimuthal angles.

Let us consider the case of the momentary radiation source when

$$g(\tau,u)=g(\tau)\delta(u). \tag{179}$$

It is the case of the particular interest. If nonstationary characteristics of the radiation field are found in the case of the momentary source, it is easy to obtain also the corresponding expression for any dependencies of function $g(\tau,u)$ from u. To do this, the expressions for characteristics of the radiation field derived for the case of the momentary source are multiplied by the function $g(\tau,u)$ and integrated over the entire interval of time during which the source is valid.

Subsequently, we supply characteristics of the radiation field created by the momentary source with index δ.

An effective way of solving problems of the non-stationary radiative transfer theory is the use of the Laplace transform in time (see, for example, Case and Zweifel 1967; Minin 1988). As a result of this conversion, we have the stationary radiative transfer equation, but the values of the absorption coefficient α and of the optical distance τ are multiplied by $1+\beta_2 s$, and the value of λ is divided by $(1+\beta_1 s)\,(1+\beta_2 s)$, where s is the parameter of the Laplace transform. The solution of the problem of nonstationary radiative transfer is obtained by using the inverse Laplace transform applied to the founded stationary characteristics of the radiation field.

Applications of the Laplace transform to the solution of nonstationary radiative transfer problems was used in works by Minin (1959, 1961, 1962, 1964, 1971, 1988). Some mathematical problems of the theory of nonstationary multiple scattering of monochromatic radiation is studied in the monograph by Wing (1962). Methods of the numerical treatment of the Laplace transform for the solution of problems of this theory are analyzed by Bellman et al. (1964).

Kolesov and Sobolev (1990) investigated the nonstationary radiative transfer in a one-dimensional homogeneous medium using the probabilistic method. The probability was determined that a photon originating at a given optical depth will emerge from the medium after multiple scattering at a given time. Knowledge this probability makes it possible to find the intensity of emergent radiation as a function of the time for arbitrary energy distribution. The result was applied to the interpretation of the flare stars luminosity variability.

6.2 *Asymptotic of the Nonstationary Radiation Field in Infinite Media*

Simple asymptotic expressions for characteristics of the non-stationary radiation field are obtained in the case when points of the medium are located at large optical distances from energy sources and scattering of light is close to conservative $(1-\lambda \ll 1)$. In this case, Minin (1988) proposed to use a simple technique for inverting of the Laplace transform.

As it is known from the theory of the Laplace transform, the value of the original at large values of the argument $(u \gg 1)$ is determined by using the expansion of the image in powers of the small parameter s. This expansion corresponds to the expansion of solutions of the stationary radiative transfer equation in powers of the small values of $1-\lambda$. Indeed, by rearranging the stationary radiative transfer equation to the corresponding nonstationary one, the value of λ is replaced by the value $\frac{\lambda}{(1+\beta_1 s)(1+\beta_2 s)}$. Therefore, taking into account the fact that $\beta_1+\beta_2=1$, we obtain $1-\lambda=s$ with accuracy to members of the higher degrees of the parameter s. Hence, when receiving the asymptotic image, it is necessary to replace the small

values of $1-\lambda$ by s in the equation solution for the stationary case, and then to apply the inverse Laplace transform.

Let us illustrate this method in the case when values of the mean intensity and the radiation flux are defined in an infinite medium illuminated by a momentary source. Let us consider three types of such sources, e.g., planar, point, and axial (see Kolesov and Kropacheva 2013, 2014). The mean intensity is defined in these cases, respectively, by expressions (149), (151), (170), and the radiation flux is given by (150), (152), and (171). Substituting s for $1-\lambda$ in these expressions, we obtain the following expressions for the Laplace transforms of these quantities:

$$\mathcal{L}[\bar{J}_\delta(\tau,u)] = \frac{lL}{8\pi}\sqrt{3-x_1}\,\frac{1}{\sqrt{s}}\mathrm{e}^{-\tau\sqrt{(3-x_1)s}}, \tag{180}$$

$$\mathcal{L}[\bar{H}_\delta(\tau,u)] = \frac{lL}{2}\mathrm{e}^{-\tau\sqrt{(3-x_1)s}} \tag{181}$$

in the case of a planar source,

$$\mathcal{L}[J_\delta(\tau,u)] = \frac{L\alpha^2}{16\pi^2\tau}(3-x_1)\mathrm{e}^{-\tau\sqrt{3-x_1)s}}, \tag{182}$$

$$\mathcal{L}[H_\delta(\tau,u)] = \frac{L\alpha^2}{4\pi\tau^2}\left[1+\tau\sqrt{(3-x_1)s}\right]\mathrm{e}^{-\tau\sqrt{(3-x_1)s}} \tag{183}$$

in the case of a point source,

$$\mathcal{L}\left[\widetilde{J}_\delta(\tau,u)\right] = \frac{lL}{16\pi^3}\sqrt{3-x_1}\,\frac{1}{\sqrt{s}}K_0\left(\tau\sqrt{(3-x_1)s}\right), \tag{184}$$

$$\mathcal{L}\left[\widetilde{H}_\delta(\tau,u)\right] = \frac{lL}{4\pi^2}K_0\left(\tau\sqrt{(3-x_1)s}\right) \tag{185}$$

in the case of an axial source. These expressions are valid for $\tau \gg 1$ and $1-\lambda \ll 1$.

The inverse Laplace transforms of these expressions lead to the following asymptotic expressions for the quantities $J_\delta(\tau,u)$ and $H_\delta(\tau,u)$:

$$\bar{J}_\delta(\tau,u) \approx \frac{lL}{8\pi\sqrt{\pi}}\sqrt{\frac{3-x_1}{u}}\exp\left(-\frac{(3-x_1)\tau^2}{4u}\right), \tag{186}$$

$$\bar{H}_\delta(\tau,u) \approx \frac{lL}{4\sqrt{\pi}}\frac{\sqrt{3-x_1}}{u\sqrt{u}}\cdot\tau\cdot\exp\left(-\frac{(3-x_1)\tau^2}{4u}\right) \tag{187}$$

for a planar source,

$$J_\delta(\tau,u) \approx \frac{L\alpha^2}{32\pi^2\sqrt{\pi}} \cdot \frac{(3-x_1)^{3/2}}{u\sqrt{u}} \exp\left(-\frac{(3-x_1)\tau^2}{4u}\right), \tag{188}$$

$$H_\delta(\tau,u) \approx \frac{L\alpha^2}{16\pi\sqrt{\pi}} \cdot \frac{(3-x_1)^{3/2}}{u^2\sqrt{u}} \cdot \tau \cdot \exp\left(-\frac{(3-x_1)\tau^2}{4u}\right) \tag{189}$$

for a point source

$$\widetilde{J}_\delta(\tau,u) \approx \frac{lL}{16\pi^3\tau} \cdot \exp\left(-\frac{(3-x_1)\tau^2}{4u}\right), \tag{190}$$

$$\widetilde{H}_\delta(\tau,u) \approx \frac{lL}{8\pi^2 u} \cdot \exp\left(-\frac{(3-x_1)\tau^2}{4u}\right) \tag{191}$$

for an axial source.

Expressions (186)–(189) give the asymptotic behavior of the mean intensity and the radiative flux under the conditions $1-\lambda \ll 1, \tau \gg 1, u > \sqrt{3-x_1}\beta_2\tau$. Expressions (190) and (191) are valid if $1-\lambda \ll 1, k\tau \gg 1, u > \beta_2\tau$.

Here we shall also give some results of studies of non-stationary radiation fields in a semi-infinite medium(see Minin 1988).

Let the surface of a semi-infinite medium be illuminated by parallel rays falling at time $u = 0$ at an angle $\arccos\xi$ to the external normal, creating the illumination of a plane perpendicular to the source equal $\pi S\delta(u)$. Let us denote by $\bar{I}_\delta(\eta,\zeta,u)$ the intensity of radiation reflected from the medium at time u in a direction forming an angle $\arccos\eta$ to the normal. Then

$$\bar{I}_\delta(\eta,\zeta,u) \approx \frac{2S}{\sqrt{3-x_1}} \frac{\zeta u_0(\eta)u_0(\zeta)}{u\sqrt{\pi u}} \quad (1-\lambda \ll 1) \tag{192}$$

The radiation intensity in the deeper layers of this medium is given by the asymptotic expression

$$\bar{I}_\delta(\tau,\eta,\zeta,u) \approx S\zeta u_0(\zeta) \frac{\tau\sqrt{3-x_1}}{2u\sqrt{\pi u}} \cdot \exp\left(-\frac{(3-x_1)\tau^2}{4u}\right) \quad (\tau \gg 1,\ 1-\lambda \ll 1) \tag{193}$$

In these expressions, $u_0(\zeta)$ is the function $u(\zeta)$ for the case $\lambda = 1$.

Let a semi-infinite medium be illuminated by momentary evenly distributed in the medium radiation sources, i.e., $g(\tau,u) = g_0\delta(u)$. In this case, the asymptotic expression for the intensity of radiation emerging from the medium has the form:

$$\bar{I}_\delta(\tau,u) \approx \frac{4g_0}{\sqrt{3-x_1}} \cdot \frac{u_0(\eta)}{\sqrt{\pi u}}. \tag{194}$$

We note that the asymptotics of the coefficients of light reflection and transmission for a planar homogeneous optically thick layer is investigated in detail in Minin (1988).

Kolesov and Sobolev (1991) considered the non-stationary radiative transfer in a one-dimensional homogeneous infinite medium with a momentary point energy source located in the origin ($\tau = 0$). The asymptotic expressions for the mean radiation intensity $J_\delta(\tau,u)$ and the radiation flux $H_\delta(\tau,u)$ were found in two limiting cases: for $t_1 \gg t_2$ (case A) and for $t_2 \gg t_1$ (case B). It was shown that

$$J_\delta(\tau,u) \approx \frac{L}{4\sqrt{\pi\lambda u}} \exp\left(-(1-\lambda)u - \frac{\tau^2}{4\lambda u}\right), \tag{195}$$

$$H_\delta(\tau,u) \approx \frac{L}{4\sqrt{\pi\lambda u}} \cdot \frac{\tau}{\lambda u} \exp\left(-(1-\lambda)u - \frac{\tau^2}{4\lambda u}\right) \tag{196}$$

in the case A under the condition $\lambda u \gg 1$, and

$$J_\delta(\tau,u) \approx \frac{L}{4}\sqrt{\frac{\lambda}{\pi u}} \exp\left(-(1-\lambda)u - \frac{\lambda\tau^2}{4u}\right), \tag{197}$$

$$H_\delta(\tau,u) \approx \frac{L\tau}{4u}\sqrt{\frac{\lambda}{\pi u}} \exp\left(-(1-\lambda)u - \frac{\lambda\tau^2}{4u}\right) \tag{198}$$

in the case B under the condition $u \gg |\tau|$.

Setting $\lambda = 1$ in (195)–(198), we have

$$J_\delta(\tau,u) \approx \frac{L}{4\sqrt{\pi u}} \exp\left(-\frac{\tau^2}{4u}\right), \tag{199}$$

$$H_\delta(\tau,u) \approx \frac{L}{4\sqrt{\pi u}} \cdot \frac{\tau}{u} \exp\left(-\frac{\tau^2}{4u}\right). \tag{200}$$

In this article, the conditions are studied under which one may use the diffusion approximation. It is shown that the replacement of the radiative transfer equation to the heat conductivity equation should give satisfactory result when $\lambda \approx 1$ but may lead to significant errors when $\lambda < 1$ (see Appendix B).

7 Concluding Remarks

Let us mention a number of other important problems of the radiative transport theory.

One of them is the polarized radiative transport problem. A brief review of this problem in the case of Rayleigh scattering was presented by Loskutov (1994). Loskutov also discussed application of this theory to astrophysical objects.

The theoretical basis for polarized light scattering was set forth, for example, by Minin (1988), Dolginov et al. (1995), Kokhanovsky (2003), Hovenier et al. (2004). The system of eigenfunctions of the polarized radiative transfer equation was investigated by Domke (1978a, b), Germogenova and Konovalov (1978), Germogenova et al. (1989), Freimanis (1990). The equation of the transfer of polarized radiation in spherically symmetric media was solved, for example, by Freimanis (1986, 2005), Bosma (1993) and Silant'ev (2006). The case of an infinite medium with cylindrically symmetric distribution of sources was considered by Freimanis (2009).

The other problem is studying the radiative transfer in inhomogeneous media. A simple model of an inhomogeneous medium is a medium consisting of an arbitrary number of homogeneous layers (see Ivanov, 1976c). If thicknesses of these layers tend to zero, we have a medium with optical characteristics depending on a spatial position of medium points. The theory of the radiative transfer in inhomogeneous plane-parallel media was set forth by Yanovitski (1997). Asymptotic characteristics of inhomogeneous radiation fields are found by Germogenova and Konovalov (1974). Wilson and Sen (1965a, b) used the method of dual P_N-approximation to study the light regime in inhomogeneous spherically symmetric media with pure isotropic monochromatic scattering. In the papers by Wilson and Sen (1965a, b), the sphere of finite radius and a spherical shell are considered under the assumption that the volume scattering coefficient is inversely proportional to the square of the distance from the center of symmetry. Radiation fields in theplane parallel to inhomogeneous media are studied by Rogovtsov and Samson (1985).

Ganapol (2008) discussed the problem of how exact, asymptotic, or numerical solutions of the radiative transfer correspond to the physical reality. Ganapol has analyzed and developed numerical techniques to control the convergence and accuracy of all the steps involved in the numerical calculations needed to achieve a benchmark-quality solution.

Appendix A: The Results of Calculation by Formula (33)

The radiation intensity $I(\tau, \mu)$ in an infinite homogeneous dusty nebula surrounding a star, was calculated according to the formula (39) with values of $\lambda = 0.9$ and $\lambda = 1$ for three phase functions: A $-$ $x(\gamma) = 1$, B $-$ $x(\gamma) = 3/4(1 + \cos^2\gamma)$, and C—$x(\gamma) = 1 + \cos\gamma + P_2(\cos\gamma)$. The calculations have been based that $\frac{L\alpha^2}{8\pi^2} = 1$. The

Table 1 Values of $I(\tau, \mu)$ for $\lambda = 0.9$

$x(\gamma)$	τ	μ				
		−1.0	−0.5	0.0	0.5	0.9
A	1.0	3.345 − 1	4.012 − 1	5.118 − 1	7.526 − 1	1.656 + 0
	2.5	6.710 − 2	8.071 − 2	1.028 − 1	1.486 − 1	2.935 − I
	5.0	9.650 − 3	1.164 − 2	1.428 − 2	2.110 − 2	3.765 − 2
	7.5	1.784 − 3	2.154 − 3	2.740 − 3	3.867 − 3	6.500 − 3
	10	3.662 − 4	4.422 − 4	5.621 − 4	7.883 − 4	1.275 − 3
	25	5.727 − 8	6.919 − 8	8.770 − 8	1.210 − 7	1.795 − 7
	50	5.721 − 14	6.912 − 14	8.747 − 14	1.197 − 13	1.723 − 15
	75	7.558 − 20	9.132 − 20	1.155 − 19	1.576 − 19	2.247 − 19
	100	1.121 − 25	1.354 − 25	1.712 − 25	2.333 − 25	3.310 − 25
B	1.0	3.858 − 1	4.174 − 1	4.975 − 1	7.218 − 1	1.692 + 0
	2.5	7.243 − 2	8.091 − 2	9.879 − 2	1.427 − 1	2.977 − 1
	5.0	1.014 − 2	1.158 − 2	1.435 − 2	2.055 − 2	3.824 − 2
	7.5	1.864 − 3	2.150 − 3	2.677 − 3	3.806 − 3	6.627 − 3
	10	3.825 − 4	4.435 − 4	5.535 − 4	7.823 − 4	1.305 − 3
	25	6.129 − 8	7.169 − 8	8.968 − 8	1.247 − 7	1.893 − 7
	50	6.455 − 14	7.570 − 14	9.469 − 14	1.306 − 13	1.919 − 13
	75	9.008 − 20	1.057 − 19	1.322 − 19	1.818 − 19	2.643 − 19
	100	1.412 − 25	1.658 − 25	2.073 − 25	2.847 − 25	4.116 − 25
C	1.0	2.402 − 1	2.384 − 1	2.941 − 1	5.187 − 1	1.605 + 0
	2.5	5.037 − 2	5.541 − 2	7.278 − 2	1.234 − 1	3.090 − 1
	5.0	8.833 − 3	1.040 − 2	1.404 − 2	2.283 − 2	4.728 − 2
	7.5	2.037 − 3	2.461 − 3	3.346 − 3	5.310 − 3	9.927 − 3
	10	5.229 − 4	6.399 − 4	8.715 − 4	1.361 − 3	2.393 − 3
	25	3.141 − 7	3.919 − 7	5.326 − 7	8.007 − 7	1.247 − 6
	50	2.936 − 12	3.682 − 12	4.990 − 12	7.386 − 12	1.106 − 11
	75	3.622 − 17	4.549 − 17	6.158 − 17	9.064 − 17	1.340 − 16
	100	5.013 − 22	6.300 − 22	8.523 − 22	1.251 − 21	1.839 − 21

calculation results are shown in Table 1 (the case $\lambda = 0.9$) and in Table 2 (the case $\lambda = 1$) for different values of the arguments τ and μ. The values of $I(\tau, \mu)$ if $\tau > 10$ in case B (the Rayleigh phase function) in Table 2 is not shown as within the accepted accuracy they coincide with the corresponding values for isotropic scattering.

When $\lambda < 1$, the radiation intensity decreases rapidly with increasing τ. This is because when $\tau \gg 1$ the values of $I(\tau, \mu)$ is proportional to $\frac{e^{-k\tau}}{\tau}$. Differences in the rate of reduction of $I(\tau, \mu)$ with increasing τ for different scattering phase functions caused mainly by the difference in the values of k (if $\lambda = 0.9$ for phase functions A, B, C the values of k are, respectively, 0.5254; 0.5232; 0.4361). When $\lambda = 1$, $k = 0$, the decrease of the values of $I(\tau, \mu)$ with increasing τ is much slower.

Table 2 Values of $I(\tau, \mu)$ for $\lambda = 1$

$x(\gamma)$	τ	μ				
		−1.0	−0.5	0.0	0.5	0.9
A	1.0	9.316 − 1	1.038 + 0	1.204 + 0	1.541 + 0	2.680 + 0
	2.5	4.367 − 1	4.945 − 1	5.493 − 1	6.460 − 1	8.839 − 1
	5.0	2.557 − 1	2.706 − 1	2.906 − 1	3.210 − 1	3.724 − 1
	7.5	1.785 − 1	1.866 − 1	1.969 − 1	2.111 − 1	2.301 − 1
	10	1.373 − 1	1.424 − 1	1.486 − 1	1.567 − 1	1.661 − 1
	25	5.777 − 2	5.878 − 2	5.990 − 2	6.116 − 2	6.231 − 2
	50	2.942 − 2	2.970 − 2	2.999 − 2	3.030 − 2	3.056 − 2
	75	1.974 − 2	1.987 − 2	2.000 − 2	2.013 − 2	2.024 − 2
	100	1.485 − 2	1.492 − 2	1.500 − 2	1.507 − 2	1.514 − 2
B	1.0	1.002 + 0	1.065 + 0	1.194 + 0	1.514 + 0	2.743 + 0
	2.5	4.668 − 1	4.963 − 1	5.448 − 1	6.399 − 1	8.947 − 1
	5.0	2.574 − 1	2.707 − 1	2.896 − 1	3.201 − 1	3.745 − 1
	7.5	1.791 − 1	1.866 − 1	1.965 − 1	2.108 − 1	2.308 − 1
	10	1.376 − 1	1.424 − 1	1.486 − 1	1.566 − 1	1.663 − 1
C	1.0	5.835 − 1	5.966 − 1	6.352 − 1	9.872 − 1	2.318 + 0
	2.5	2.791 − 1	2.986 − 1	3.395 − 1	4.328 − 1	7.084 − 1
	5.0	1.596 − 1	1.708 − 1	1.883 − 1	2.189 − 1	2.772 − 1
	7.5	1.132 − 1	1.199 − 1	1.294 − 1	1.439 − 1	1.649 − 1
	10	8.793 − 2	9.240 − 2	9.826 − 2	1.065 − 1	1.167 − 1
	25	3.781 − 2	3.878 − 2	3.988 − 2	4.116 − 2	4.235 − 2
	50	1 943 − 2	1.970 − 2	1.998 − 2	2.030 − 2	2.056 − 2
	75	1.308 − 2	1.320 − 2	1.333 − 2	1.347 − 2	1.358 − 2
	100	9.853 − 3	9.924 − 3	9.998 − 3	1.007 − 2	1.014 − 2

The values $I(\tau, \mu)$ increase which increasing μ. It is especially noticeable when $\lambda \approx 1$. The radiation intensity becomes infinite in the direction of radiation entered a given point of medium directly from the star (when $\mu = 1$).

Appendix B: The Exactness of Approximate Expressions for the Mean Radiation Intensity and the Radiation Flux

The nonstationary radiation field in a one-dimensional homogeneous infinite medium with a point energy source is calculated in the cases A ($t_1 \gg t_2$) and B ($t_2 \gg t_1$). Obtained exact values for the mean radiation intensity $J_A(\tau, u)$, $J_B(\tau, u)$ and the radiation flux $H_A(\tau, u)$, $H_B(\tau, u)$ are compared with corresponding values $J_D(\tau, u)$ and $H_D(\tau, u)$ founded in the diffusion approximation. Results of the comparison are shown in Tables 3, 4, and 5.

Table 3 Ratios of J_A/J_D and H_A/H_D for $\lambda = 1$

τ	1		5		10	
u	J_A/J_D	H_A/H_D	J_A/J_D	H_A/H_D	J_A/J_D	H_A/H_D
I	0.801	0.98	19.3	5.80	5.30×10^7	8.57×10^6
2	0.916	1.57	2.01	1.017	7.78×10^2	2.24×10^2
3	0.977	1.76	1.18	0.779	2.92×10	1.15×10
4	0.995	1.77	0.997	0.768	6.92	3.33
5	1.018	1.69	0.939	0.802	3.27	1.82
6	1.022	1.59	0.922	0.845	2.12	1.32
7	1.024	1.50	0.921	0.887	1.63	1.104
8	1.023	1.43	0.926	0.923	1.38	0.997
9	1.022	1.36	0.934	0.953	1.23	0.942
10	1.021	1.32	0.941	0.976	1.14	0.914
15	1.015	1.19	0.970	1.038	0.993	0.906
20	1.012	1.13	0.984	1.052	0.971	0.939
30	1.008	1.083	0.995	1.050	0.977	0.983
40	1.006	1.060	0.999	1.043	0.985	1.001
50	1.005	1.047	1.000	1.037	0.990	1.009
60	1.004	1.039	1.001	1.032	0.993	1.012
80	1.003	1.029	1.001	1.025	0.996	1.014
100	1.002	1.023	1.001	1.021	0.998	1.014

Consider first the case A. In the case of pure scattering (i.e., $\lambda = 1$), asymptotic expression (197) and (198) when $u \gg 1$ just coincide with formulas (199) and (200) for $J_D(\tau, u)$ and $H_D(\tau, u)$. Exact and approximate values of J and H are quite close to each other (see Table 3). So we can assume that in the case A when $\lambda = 1$, the application of diffusion approximation will not lead to large errors.

Another situation that occurs is in the presence of true absorption, i.e., when $\lambda < 1$. The exact values of $J_A(\tau, u)$ and $H_A(\tau, u)$ may differ significantly from their asymptotic values. Asymptotic expressions (197) and (198) also differ from expressions (199) and (200). The radio J_A/J_D may differ significantly from unity.

Addressing the case B, we first of all note that in this case always $u \gg \tau$, e.g. $J_B(\tau, u) = 0$ and $H_B(\tau, u) = 0$ when $u < \tau$. But the diffusion equation does not take into account the finite speed of light. Taking into account only $u > \tau$, approximately the same results, as in case A, can be obtained by exact and approximate expressions. When $\lambda = 1$, an asymptotic expressions for $J_B(\tau, u)$ and $H_B(\tau, u)$ coincides with expressions (199) and (200) for $J_D(\tau, u)$ and $H_D(\tau, u)$.

The ratio J_B/J_D and H_B/H_D greatly depends on λ (see Tables 4 and 5).

From the above it is concluded that the replacement of the radiation transfer equation to the diffusion equation should give satisfactory results when $\lambda \approx 1$ but may lead to significant errors when $\lambda < 1$.

Table 4 Values of $J_A(\tau, u)$,$J_D(\tau, u)$, $J_B(\tau, u)$, for $\lambda = 0.5$

τ	1			5			10		
u	$J_A(\tau, u)$	$J_D(\tau, u)$	$J_B(\tau, u)$	$J_A(\tau, u)$	$J_D(\tau, u)$	$J_B(\tau, u)$	$J_A(\tau, u)$	$J_D(\tau, u)$	$J_B(\tau, u)$
I	5.50 − 2	6.66 − 2	3.32 − 2	2.08 − 3	1.65 − 4	0	2.84 − 5	1.19 − 12	0
2	3.24 − 2	3.24 − 2	1.82 − 2	1.93 − 3	1.61 − 3	0	3.82 − 5	1.37 − 7	0
3	1.88 − 2	1.67 − 2	1.01 − 2	1.57 − 3	2.26 − 3	0	4.11 − 5	4.37 − 6	0
4	1.09 − 2	8.96 − 3	5.63 − 3	1.19 − 3	2.00 − 3	0	3.91 − 5	1.84 − 5	0
5	6.30 − 3	4.92 − 3	3.18 − 3	8.58 − 4	1.48 − 3	2.39 − 3	3.43 − 5	3.49 − 5	0
6	3.64 − 3	2.75 − 3	1.82 − 3	5.98 − 4	1.01 − 3	1.39 − 3	2.84 − 5	4.44 − 5	0
7	2.11 − 3	1.55 − 3	1.04 − 3	4.06 − 4	6.59 − 4	8.07 − 4	2.25 − 5	4.53 − 5	0
8	1.22 − 3	8.85 − 4	5.94 − 4	2.70 − 4	4.18 − 4	4.72 − 4	1.73 − 5	4.01 − 5	0
9	7.08 − 4	5.08 − 4	3.43 − 4	1.77 − 4	2.61 − 4	2.77 − 4	1.29 − 5	3.25 − 5	0
10	4.11 − 4	2.93 − 4	1.99 − 4	1.14 − 4	1.61 − 4	1.63 − 4	9.37 − 6	2.47 − 5	7.78 − 5
15	2.82 − 5	1.98 − 5	1.36 − 5	1.14 − 5	1.33 − 5	1.17 − 5	1.51 − 6	3.80 − 6	6.95 − 6
20	2.02 − 6	1.41 − 6	9.81 − 7	1.02 − 6	1.05 − 6	8.68 − 7	1.90 − 7	4.10 − 7	5.79 − 7
30	1.11 − 8	7.81 − 9	5.46 − 9	7.12 − 9	6.40 − 9	5.00 − 9	2.09 − 9	3.42 − 9	3.78 − 9
40	6.50 − 11	4.57 − 11	3.20 − 11	4.68 − 11	3.93 − 11	2.99 − 11	1.81 − 11	2.46 − 11	2.41 − 11
50	3.92 − 13	2.76 − 13	1.93 − 13	3.02 − 13	2.44 − 13	1.83 − 13	1.40 − 13	1.68 − 13	1.54 − 13
60	2.41 − 15	1.70 − 15	1.19 − 15	1.95 − 15	1.54 − 15	1.14 − 15	1.02 − 15	1.12 − 15	9.82 − 16
80	9.47 − 20	6.68 − 20	4.70 − 20	8.10 − 20	6.20 − 20	4.54 − 20	4.99 − 20	4.90 − 20	4.06 − 20
100	3.85 − 24	2.71 − 24	1.91 − 24	3.40 − 24	2.56 − 24	1.86 − 24	2.32 − 24	2.12 − 24	1.70 − 24

Table 5 Values of $H_A(\tau, u)$, $H_D(\tau, u)$, $H_B(\tau, u)$, for $\lambda = 0.5$

τ	1			5			10		
u	$H_A(\tau, u)$	$H_D(\tau, u)$	$H_B(\tau, u)$	$H_A(\tau, u)$	$H_D(\tau, u)$	$H_B(\tau, u)$	$H_A(\tau, u)$	$H_D(\tau, u)$	$H_B(\tau, u)$
I	8.69 − 2	6.66 − 2	7.38 − 3	3.50 − 3	8.26 − 4	0	4.96 − 5	1.19 − 11	0
2	4.12 − 2	1.62 − 2	3.57 − 3	2.90 − 3	4.03 − 3	0	6.18 − 5	6.84 − 7	0
3	1.96 − 2	5.57 − 3	1.75 − 3	2.16 − 3	3.77 − 3	0	6.26 − 5	1.46 − 5	0
4	9.38 − 3	2.24 − 3	8.73 − 4	1.51 − 3	2.50 − 3	0	5.64 − 5	4.61 − 5	0
5	4.52 − 3	9.85 − 4	4.41 − 4	1.01 − 3	1.48 − 3	1.84 − 3	4.71 − 5	6.98 − 5	0
6	2.20 − 3	4.58 − 4	2.26 − 4	6.56 − 4	8.43 − 4	9.45 − 4	3.73 − 5	7.41 − 5	0
7	1.08 − 3	2.22 − 4	1.17 − 4	4.16 − 4	4.71 − 4	4.92 − 4	2.83 − 5	6.47 − 5	0
8	5.37 − 4	1.11 − 4	6.12 − 5	2.60 − 4	2.61 − 4	2.59 − 4	2.08 − 5	5.02 − 5	0
9	2.69 − 4	5.64 − 5	3.24 − 5	1.60 − 4	1.45 − 4	1.38 − 4	1.49 − 5	3.61 − 5	0
10	1.37 − 4	2.93 − 5	1.73 − 5	9.77 − 5	8.04 − 5	7.40 − 5	1.04 − 5	2.47 − 5	8.64 − 5
15	5.37 − 6	1.32 − 6	8.39 − 7	7.47 − 6	4.43 − 6	3.68 − 6	1.41 − 6	2.54 − 6	4.78 − 6
20	2.60 − 7	7.07 − 8	4.63 − 8	5.29 − 7	2.62 − 7	2.08 − 7	1.52 − 7	2.05 − 7	2.91 − 7
30	8.66 − 10	2.60 − 10	1.75 − 10	2.54 − 9	1.07 − 9	8.08 − 10	1.27 − 9	1.14 − 9	1.25 − 9
40	3.63 − 12	1.14 − 12	7.79 − 13	1.25 − 11	4.92 − 12	3.65 − 12	8.74 − 12	6.15 − 12	5.60 − 12
50	1.71 − 14	5.51 − 15	3.79 − 15	6.43 − 14	2.44 − 14	1.80 − 14	5.57 − 14	3.36 − 14	3.04 − 14
60	8.62 − 17	2.83 − 17	1.95 − 17	3.42 − 16	1.28 − 16	9.33 − 17	3.43 − 16	1.87 − 16	1.62 − 19
80	2.50 − 21	8.35 − 22	5.80 − 22	1.06 − 20	3.87 − 21	2.80 − 21	1.27 − 20	6.13 − 21	5.04 − 21
100	8.02 − 26	2.71 − 26	1.89 − 26	3.52 − 25	1.28 − 25	9.20 − 26	4.71 − 25	2.12 − 25	1.63 − 25

References

Abramowitz M, Stegun A (eds) (1964) Handbook of mathematical functions with formulas, graphs and mathematical tables. National Bureau of Standards, New York

Ambartsumyan VA (1943) On the problem of diffuse light reflection by a turbid medium. Dokl Akad Nauk SSSR 38(8):257–261

Ambartsumyan VA (1945) Point source of light in a turbid medium. Bull Yerevan Astron Obs 6:3–9

Ambartsumyan VA (1960) Scientific works, vol 1. Izd. Akad. Nauk Armen. SSR, Yerevan

Ambartsumyan VA (1998) In: Ambartsumyan RV (ed) A life in astrophysics: selected papers of Viktor Ambartsumyan. Allerton Press, New York

Apruzese JP (1981) Direct solution of the equation of transfer using frequency and angle-averaged photon-escape probabilities for spherical and cylindrical geometries. J Quant Spectroscop Radiat Transfer 25(5):419–425

Avery LW, House LL, Skumanich A (1969) Radiative transport in finite homogeneous cylinders by the MonteCarlo technique. J Quant Spectrosc Radiat Transfer 9(4):519–531

Bailey PB (1964) A regorous derivation of some invariant imbedding equations of transport theory. J Math Anal Appl 8(1):144–169

Bailey PB, Wing GM (1964) A correction to some invariant imbedding equation of transport theory obtained by particle counting. J Math Anal Appl 8(1):170–174

Barkov VI (1960) Radiative transfer in a homogeneous medium for non-conservative case. Optika i Spectroscopiya 9(3):376–385

Barkov VI (1963) Radiative transfer in a homogeneous sphere with a central point source. Optika i Spectroscopiya 14(4):537–544

Bell D, Glasstone S (1970) Nuclear reactor theory. Van Nostre and Reinhold Co., Princeton

Bellman R, Kalaba R (1957) On the principle of invariant imbedding and diffuse refection from cylindrical regions. Proc Nat Acad Sci USA 43(6):514–517

Bellman R, Kalaba R, Wing GM (1959) Invariant imbedding and neutron transport theory. IV. Generalized transport theory. J Math Mech 8(4):575–584

Bellman R, Kalaba R, Wing GM (1960) Invariant imbedding and mathematical physics. I. Particle processes. J Math Phys 1(4):280–308

Bellman R, Kagiwada H, Kalaba R, Prestrud MC (1964) Invariant imbedding and time-dependent transport processes. Elsevier Publishing Co., New York

Bellman R, Kalaba R (1965) On the new approach to the numerical solution of a class of particle differential integral equation of transport theory. Proc Nat Acad Sci USA 54(5):1293–1296

Bellman R, Kagiwada H, Kalaba R (1966) Invariant imbedding and radiative transfer in spherical shells. J Comput Phys 1(2):245–256

Bellman R, Kagiwada H, Kalaba R, Ueno S (1968) Diffuse transmission of light from a central source through an inhomogeneous spherical shell with isotropic scattering. J Math Phys 9 (6):909–912

Bellman R (1968) A new technique from the numerical solution of Fredholm integral equations. Computing 3(2):131–138

Bosma PB (1993) The intensity and state polarization of light scattering in a spherical shell. Astron Astrophys 279(2):572–576

Brand PWJL (1979) Scattering from a cylindrical dust cloud in an isotropic radiation field. Astron Astrophys 71(1–2):47–50

Busbridge IW (1960) The mathematics of radiative transfer. Cambridge University Press, Cambridge

Case KM, Hoffman F, Placzek G (1953) Introduction to the theory of neutron diffusion, vol 1. Government Printing Office, Washington

Case KM (1960) Elementary solutions of the transport equation and their applications. Ann Phys 9 (1):1–23

Case KM, Zweifel PF (1967) Linear transport theory. Addison-Wesley Pub. Co, Reading

Case KM, Zalazny R, Kanal M (1970) Spherically symmetric boundary-value problems in one-speed transport theory. J Math Phys 11(1):223–239
Cassinelli JP, Hummer DG (1971) Radiative transfer in spherically symmetric systems. II. The non-conservative case and linearly polarized radiation. Mon Notic Roy Astron Soc 154(1):9–21
Casti J, Kalaba R (1973) Imbedding methods in applied mathematics. Addison-Wesley Pub. Co., Reading
Chandrasekhar S (1934a) The radiative equilibrium of extended stellar atmospheres. Mon Notic Roy Astron Soc 94(5):444–458
Chandrasekhar S (1934b) An analysis of the problems of the stellar atmospheres. Astron J Soviet Union 11(6):550–596
Chandrasekhar S (1935) On the effective temperatures of extended photospheres. Proc Cambridge Phil Soc 31(3):390–393
Chandrasekhar S (1945) On the radiative equilibrium of a stellar atmosphere. V. Astrophys J 101 (1):95–107
Chandrasekhar S (1960) Radiative transfer. Oxford University Press, Oxford
Chapman RD (1966) Radiative transfer in extended stellar atmospheres. Astrophys J 143(1):61–74
Chou YS, Tien CL (1968) A modified moment method for radiative transfer in non-planar systems. J Quant Spectrosc Radiat Transfer 8(3):919–935
Crosbie AL, Khalil HK (1972) Radiative transfer in a gray isothermal spherical layer. J Quant Spectrosc Radiat Transfer 12(10):1465–1486
Davison B (1957a) Spherical-harmonics method for neutron-transport problems in cylindrical geometry. Ganad J Phys 35(5):576–593
Davison B (1957b) Neutron transport theory. Clarendon Press, Oxford
Davydov VI, Shikhov SB (1973) Analytical solution of the system of the Carlson S_n-equations for spherical geometry. Physics of nuclear reactors, vol 3. Atomizdat, Moscow, pp 22–28
Davydov VI, Shikhov SB (1975) Analytical solution methods for the gas kinetic equation of neutron transport. Atomizdat, Moscow
DeBar RB (1967) Difference equations for the Legendre polinomial representation of the transport equation. J Comput Phys 2(2):197–205
Dolginov AZ, Gnedin YuN, Silant'ev NA (1995) Propagation and polarization of radiation in cosmic media. Gordon and Breach, Amsterdam
Dolin LS, Luchinin AG (1971) Relationship between radiation fields of plane and point isotropic sources in a turbid medium. Izv Akad Nauk SSSR Fiz Atmos Okeana 7(10):1103–1106
Dolin LS (1981) On the methods of calculation of brightness for large optical distance from the radiation sources. Izv Akad Nauk SSSR Fiz Atmos Okeana 17(1):102–105
Domke H, Ivanov VV (1975) Asymptotic properties of the Green's function of the equation of transport of polarized radiation. Astron Zh 52(5):1034–1037
Domke H (1978a) Linear Fredholm equation for radiative transfer problems in finite plane-parallel media. I. Imbedding in an infinite medium. Astron Nachr 299:95–102
Domke H (1978b) Linear Fredholm equation for radiative transfer problems in finite plane-parallelmedia. II. Imbedding in a semi-infinite medium. Astron Nachr 299:87–93
Domke H (1983) Biorthogonality and radiative transfer in finite slab atmospheres. J Quant Spectrosc Radiat Transfer 30(2):119–129
Dorschner J (1971) Radiative transfer in spherically symmetric dust nebulae. Astron Nachr 292(5–6):225–229
Drawbaugh D, Noderer L (1959) The double spherical harmonics method for cylinders and spheres. Nucl Sci Eng 6(1):79–81
Dubi A, Horowitz YS (1978) A semi-analytic solution to the transport equation in spherical geometry. Nucl Sci Eng 66(1):118–121
Eddington AS (1926) The internal constitution of the stars. Cambridge University Press, Cambridge
Erdman RC, Siewert GE (1968) Green's functions for the one-speed transport equation in spherical geometry. J Math Phys 9(1):81–89

Ermakov SM, Mikhailov GA (1982) Statistical modelling. Izd. Nauka, Moscow
Ershov YI, Shikhov SB (1972) Integral transforms as a means of solving boundary value problems in neutron transport theory. USSR Comp Matem Matem Fiz 12(3):87–107
Ershov YI, Shikhov SB (1973) The solution of transfer theory problems in spherical geometry. USSR Comp Matem Matem Fiz 13(6):202–212
Ershov YI, Shikhov SB (1977) Methods for solving boundary value problems in transport theory. Atomizdat, Moscow
Ershov YI, Shikhov SB (1985) Mathematical foundation of transport theory, vol 1, 2. Energoatomizdat, Moscow
Freimanis J (1986) A point source in an infinite medium. Nauch Inf Astrosoveta 61:95–109
Freimanis J (1990) On the completeness of system of eigenfunctions and adjoin functions of radiative transfer equation of polarized radiation. Invest Sun Red Stars 32:20–116 (Riga)
Freimanis J (2005) On Green's functions for spherically symmetric problems of transferof polarized radiation. J Quant Spectosc Radiat Transfer 96:451–472
Freimanis J (2009) On Green's functions for cylindrically symmetric fields of polarized radiation. J Quant Spectosc Radiat Transfer 110:1307–1334
Gakhov FD (1966) Boundary value problems. Pregamon Press, Oxford
Gakhov FD, Cherskij UI (1978) Equations of convolution type. Izd. Nauka, Moscow
Ganapol BD (2008) Analytical benchmarks for nuclear engineering applications: case studies in neutron transport theory. Nuclear Energy Agency, Paris
Germogenova TA (1966) Diffusion of the radiation in a spherical shell surrounded a point source. Astrophysics 2(3):125–131
Germogenova TA (1971) Numerical methods for solving boundary-value problems for the transport equation. Theoretical and applied problems of light scattering. Izd. Nauka i Tekhnika, Minsk, pp 29–42
Germogenova TA (1972) On regular solutions of the transport characteristic equation, vol 21. Preprint of Inst. Prikl. Matem., Akad Nauk SSSR, Moscow
Germogenova TA (1974) On the discrete spectrum of the transport characteristic equation. USSR Comp Matem Matem Fiz 14(6):149–165
Germogenova TA, Konovalov NV (1974) Asymptotic characteristics of the solution of the transport equation in the inhomogeneous layer problem. USSR Comp Matem Matem Fiz 14 (4):107–125
Germogenova TA, Konovalov NV (1978) The spectrum of the characteristic equation in the theory of transfer with polarization taken into account, vol 62. Preprint of Inst. Prikl. Matem., Akad. Nauk SSSR, Moscow
Germogenova TA (1986) Local properties of solutions of the transport equation. Izd. Nauka, Moscow
Germogenova TA, Konovalov NV, Kuz'mina MG (1989) 'Foundations of the mathematical theory of polarizedradiation transfer (rigorous results). In: Mnatsakanian MA, Pikichian OV (eds) The invariance principle and its applications. Izd. Akad. Nauk Armen. SSR, Yerevan
Germogenova TA, Pavelyeva EB (1989) The characteristic equation in radiation transport problems in elongated cylindrical regions. USSR Comp Matem Matem Fiz 29(4):158–170
Gorelov VP, Yuferev VI (1971) The solution of the one-speed neutron transport equation in multilayer plane and spherically symmetric systems. USSR Comp Matem Matem Fiz 11(1):171–181
Gorelov VP, Ilyin VI, Yuferev VI (1972) Solution of some problems of neutron transport theory by the method of generalized eigenfunctions. USSR Comp Matem Matem Fiz 12(5):200–224
Grachov SI (1978) On a problem of radiative diffusion in moving media, II. Spherical symmetry, Vest Lenengrad Univ 1:129–135
Grachov SI (1994) Radiative transfer in moving astrophysical media. Trudy Astron Obs St Petersburg Gos Univ 44:203–235
Grant IP, Hunt GE (1969a) Discrete space theory of radiative transfer. I. Fundamentals. Proc Roy Soc London Ser A 313(1513):183–197

Grant IP, Hunt GE (1969b) Discrete space theory of radiative transfer. II. Stability and non-negativity. Proc Roy Soc London Ser A 313(1513):199–216
Greenspan H, Kelber CN, Okrent D (eds) (1968) Computing Methods in Reactor Physics. Gordon and Breach Science Publishing, New York
Grinin VP (1994) Non-stationary radiative transfer theory. Trudy Astron Obs St Petersburg Gos Univ 44:236–249
Gritton EC, Leonard A (1970) Exact solution to the radiation heat transport equation in gaseous media using singular integral equation theory. J Quant Spectrosc Radiat Transfer 10(10):1095–1118
Gruschinske J, Ueno S (1971) Bellman's new approach to the numerical solution of Fredholm integral equations with positive kernels. J Quant Spectrosc Radiat Transfer 11(6):641–646
Heaslet MA, Warming RP (1965) Application of invariance principles to a radiative transfer problem in a homogeneous spherical medium. J Quant Spectrosc Radiat Transfer 5(5):669–682
Heaslet MA, Warming RP (1966) Theoretical predictions of radiative transfer in a homogeneous cylindrical medium. J Quant Spectrosc Radiat Transfer 6(6):751–774
Hovenier JW, van der Mee C, Domke H (2004) Transfer of polarized light in planetary atmospheres. Basic concepts and practical methods. Kluwer Academic Publisher, Chichester
Hummer DG, Rybicki GB (1971) Radiative transfer in spherically symmetric systems. The conservative gray case. Mon Notic Roy Astron Soc 152(1):1–19
Hummer DG, Kunasz CV, Kunasz PB (1973) Numerical evaluation of the radiative transfer problems in spherical geometries. Comp Phys Comm 6(1):38–57
Hund GE (1968) Radiative transfer in a homogeneous cylindrical atmosphere. SIAM J Math 16 (6):1255–1265
Ivanov VV (1969) Mean number of photon scattering in a homogeneous optically thick sphere. Stars, nebulae, galaxies. Izd. Akad. Nauk Armen. SSR, Yerevan, pp 27–39
Ivanov VV (1973) Transfer of radiation in spectral lines, vol 385. National Bureau of Standards, Special Publication, Washington
Ivanov VV (1976a) Radiative transfer in an infinite atmosphere. I. Astrophysics 12(2):158–168
Ivanov VV (1976b) Radiative transfer in an infinite atmosphere. II. Astrophysics 12(4):369–379
Ivanov VV (1976c) Radiative transfer in a multilayer optically thick atmosphere. I, II. Trudy Astron Obs Leningrad Gos Univ 32:3–39
Ivanov VV, Volkov EV (1979) Radiative transfer. Principles of invariance for the Green's functions. Trudy Astron Obs Leningrad Gos Univ 35:3–30
Ivanov VV (1991) 100 years the integral equation of radiative transfer. In: Ivanov AP (ed) Scattering and absorption of light in natural and artificial turbid media. Izd Inst. Fiz. AN Belarus, Minsk, pp 10–36
Ivanov VV (1994) Making of radiative transfer theory. Trudy Astron Obs St Petersburg Gos Univ 44:6–29
Inönü E (1970) Orthogonality of a set of polynomials encountered in neutron-transport and radiative-transport theories. J Math Phys 11(2):568–577
Kalaba R, Ruspini EH (1971) Invariant imbedding and radiative transfer in a homogeneous spherical medium. J Quant Spectrosc Radiat Transfer 11(7):1063–1074
Kho TH, Sen KK (1971) On combined operational method for transfer problems in homogeneous spherical media. Astrophys Space Sci 14(2):223–242
Kho TH, Sen KK (1972) On combined operational method for transfer problems in homogeneous cylindrical media. Astrophys Space Sci 16(1):151–166
Kokhanovsky AA (2003) Polarization Optics of Random Media. Springer-Praxis, Chichester
Kolesov AK (1983) Point source in an absorbing and anisotropically scattering infinite uniform medium. Sov Phys Dokl 28:700–703
Kolesov AK (1984a) The Green's functions for equation of radiative transfer in an infinite homogeneous medium with a radial distribution of the sources. Astrophysics 20(1):86–95
Kolesov AK (1984b) The asymptotic radiation regime in the outer layers of a homogeneous sphere of large optical radius. Astrophysics 21(2):519–527

Kolesov AK (1985a) Asymptotic relations in the theory of radiative transfer in a sphere and a spherical shell. Astrophysics 22(1):107–114
Kolesov AK (1985b) Radiation field in media with spherical symmetry. Astrophysics 22(3):339–346
Kolesov AK (1985c) On the light regime in a homogeneous spherical envelope. Vest Leningrad Univ 8:73–80
Kolesov AK (1986) Radiative transfer in media with cylindrical symmetry. Sov Phys Dokl 31:241–244
Kolesov AK, Perov VJu (1987) Radiation field in an infinite dust nebula illuminated by a star. Astrophysics 26(2):147–154
Kolesov AK (1990) The radiation field in an infinite medium far from a point source. Vest Leningrad Univ 22:62–67
Kolesov AK, Sobolev VV (1990) Nonstationary radiative transfer in stellar atmospheres. Sov Astron 34(2):179–188
Kolesov AK, Sobolev VV (1991) On non-stationary radiative transfer. Trudy Astron Obs Leningrad Univ 43:5–27
Kolesov AK (1994) Light scattering in spherically symmetric media. Trudy Astron Obs St Petersburg Gos Univ 44:114–130
Kolesov AK, Kropacheva NY (2013) Some asymptotic expressions in the theory of non-stationary transfer of radiation. Vest St Petersburg Univ 4(1):127–133
Kolesov AK, Kropacheva NY (2014) On asymptotic light regime in an infinite medium far from a linear source of energy. Vest St Petersburg Univ 2(1):152–158
Konovalov NN (1974) Asymptotic properties of one-speed radiative transfer solution in homogeneous plane layer. Problems of azimuthal dependence, vol 65. Preprint of the Inst. Prikl. Matem., Akad. Nauk SSSR, Moscow
Kosirev NA (1934) Radiative equilibrium of the extended photosphere. Mon Notic Roy Astron Soc 94(5):430–443
Kourganoff V (1952) Basic methods in transfer problems. Clarendon Press, Oxford
Kunasz PB, Hummer DG (1974) Radiative transfer in spherically symmetric systems. III. Fundamentals of line formation. Mon Notic Roy Astron Soc 166(1):19–55
Kuznetsov ES (1951) Ratiative equilibrium of a gaseous envelope surrounding an absolutely black sphere. Isv Akad Nauk SSSR Ser Geofiz 15(3):69–93
Kuznetsov ES (1962) Temperature distribution in an infinite cylinder and in a sphere in a state of non-monochromatic radiation equilibrium. USSR Comp Matem Matem Fiz 2(2):230–245
Kurchakov AV (1970) An airglow of a homogeneous sphere in the field of the isotropic radiation. Sov Astron A J 47(6):1339–1340
Laletin NI (1966) Elementary solutions for neutron transport equations in problems with cylindrical and spherical symmetry. Atom Energy 20(6):585–586
Laletin NI (1969a) Green's functions for the neutron transport equation for problems with cylindrical and spherical symmetry. Atom Energy 26(4):370–371
Laletin NI (1969b) Methods of surface psevdosources for solving the transport equation (G_N-approximations). In: Marchuk GI (ed) Computional methods in the transport theory. Atomizdat, Moscow, pp 228–249
Laletin NI (1974a) Elementary solutions of the one-speed neutron transport equation. In: Shevelev VY (ed) Methods of calculation of thermal neutron fields in the reactors lattices. Atomizdat, Moscow, pp 155–186
Laletin NI (1974b) Method of surface psevdosources for solving the neutron transport equation (G_N-approximations). In: Shevelev VY (ed) Methods of calculation of thermal neutron fields in the reactors lattices. Atomizdat, Moscow, pp 187–215
Leonard A, Mullikin TW (1964) Green's functions for one-speed neutron transport in an one-dimensional slab and sphere. Proc Nat Acad Sci USA 52(3):683–688
Leong TK, Sen KK (1968) Probabilistic model for transfer problem in an externally illuminated spherical shell atmosphere with a perfectly absorbing core. Ann Astrophys 31(5):467–474

Leong TK, Sen KK (1969) Probabilistic model for radiative transfer problems for spherical shell medium. Publ Astron Soc Jpn 21(2):167–175
Leong TK, Sen KK (1970) Probabilistic model for radiative transfer problems for inhomogeneous infinite cylindrical shell medium. Publ Astron Soc Jpn 22(1):57–74
Leong TK, Sen KK (1971a) An integral equation method for transfer problems in spherical shell media. Publ Astron Soc Jpn 23(1):99–115
Leong TK, Sen KK (1971b) An integral equation method for transfer problems in infinite cylindrical shell media. Publ Astron Soc Jpn 23(2):247–262
Leong TK, Sen KK (1972) A probabilistic model for time-dependent transfer problems in spherical shell media. Mon Notic Roy Astron Soc 160(1):21–36
Leong TK (1972) Probabilistic model for time-dependent transfer problems in finite cylindrical shell medium. Astrophys Space Sci 19(2):369–385
Leong TK, Sen KK (1973a) Probabilistic model for radiative transfer problems in spherical shell media with complete redistribution in frequency. Mon Notic Roy Astron Soc 163(1):105–116
Leong TK, Sen KK (1973b) Probabilistic model for radiative transfer problems in cylindrical shell media with complete redistribution in frequency. Astrophys Space Sci 21(1):59–72
Leung CM (1975) Radiation transport in dense interstellar dust clouds. I. Grain temperature. Astrophys J 199(2):340–360
Loskutov VM (1974) The albedo of a homogeneous sphere. Vest Leningrad Univ 19:132–135
Loskutov VM (1994) 'Transfer of polarized radiation: Rayleigh scattering. Trudy Astron Obs St Petersburg Gos Univ 44:154–171
Marchuk GI (1961) Methods of calculations of nuclear reactors. Gosatomizdat, Moscow
Marchuk GI, Mikhailov GA (1967) Results of the solution of some problems in atmospheric optics by a Monte Carlo method. Atmos Oceanic Phys 3:167–171
Marchuk GI, Lebedev VI (1981) Numerical methods in the of neutron transport theory, 2nd edn. Atomizdat, Moscow
Marshak RE (1947) Note on the spherical harmonic method as applied to the Milne problem for a sphere. Phys Rev 71(7):443–446
Maslennikov MV (1969) The Milne problem with anisotropic scattering. American Mathematical Society, Providence, Rhode Island
McCrea WH (1928) A note on Dr. P.A.Taylor's paper 'The equilibrium of the calcium chromosphere. Mon Notic Roy Astron Soc 88(7):729–740
Meyer JF, Jacobs AM (1970) Point source Green's functions for neutral particle transport. Nucl Sci Eng 40(2):239–245
Mika JR (1961) Neutron transport with anisotropic scattering. Nucl Sci Eng 11(4):415–427
Mikhailov GA (1987) Optimization of weighting Monte Carlo methods. Izd. Nauka, Moscow
Milne EA (1921) Radiative equilibrium in the outer layers of a star: the temperature distribution and the law of darkening. Mon Notic Roy Astron Soc 81:361–388
Minin IN (1959) Non-stationary problems of the radiative transfer theory. Vest Leningrad Univ 13:137–141
Minin IN (1961) Diffuse reflection from a semi-infinite medium for anisotropic scattering, I. Vest Leningrad Univ 1:133–143
Minin IN (1962) On the theory of non-stationary diffusion of radiation. Vest Leningrad Univ 19:124–132
Minin IN (1964) On nonstationary glow of a semi-infinite medium. Sov Phys Dokl 9:114–117
Minin IN (1966) Diffusion of radiation in a plane layer for anisotropic scattering. Astron Zh 43 (6):1244–1260
Minin IN (1971) Nonsteady-state emission of a one-dimensional medium of finite optical thickness. Sov Astron A J 15:260–265
Minin IN (1988) Theory of radiative transfer in planetary atmospheres. Izd. Nauka, Moscow
Mishchenko MI, Dlugach JM, Yanovitskij EG, Zakharova NI (1999) Bidirectional reflectance of flat, optically thick particulate layers: an efficient radiative transfer solution and applications to snow and soil surfaces. J Quant Spectrosc Raiat Transfer 63(2–6):409–432

Mishchenko MI, Dlugach JM, Chondhary J, Zakharova NI (2015) Polarized bidirectional reflectance of optically thick sparse particulate layers: an efficient numerically exact radiative transfer solution. J Quant Spectrosc Radiat Transfer 156:97–108
Mnatsakanian MA (1976) Quasi-asymptotic solutions of the problem of radiative transfer in a finite-thickness layer. II—nonconservative scattering. Astrofizika 12(3):451–473
Muskhelishvili NI (1962) Singular integral equations. Fizmatgiz, Moscow
Nagirner DI (1965) On the light scattering in a spherical nebulae of large optical thickness. Trudy Astron Obs Leningrad Gos Univ 22:66–71
Nagirner DI (1972) Scattering of resonance radiation in a sphere. Astrophysics 8(3):210–219
Nagirner DI (1974) Theory of nonstationary transfer of radiation. Astrophysics 10(3):274–289
Nagirner DI (1986) Transport of radiation in a layer, sphere, and cylinder. Sov Phys Dokl 31:577–580
Nagirner DI (1994a) Radiative transfer in a cylinder. I. The resolvent of the basic integral equation. Astrophysics 37(1):69–78
Nagirner DI (1994b) Radiative transfer in a cylinder. II. Special problems. Asymptotics. Astrophysics 37(4):362–371
Nagirner DI (1995) Radiative transfer in a cylinder. III. Spectrum of the basic integral equation. Astrofizika 38(1):39–47
Nonnenmacher TF (1967) Elementary solutions of the monoenergetic Boltzmann equation in spherical geometry. Nukleonik 10(5):275–277
Özisik MN, Menning J, Hälg W (1975) Half-range moment method for solution of the transport equation in a spherically symmetric geometry. J Quant Spectrosc Radiat Transfer 15(12):1101–1106
Pavelyeva EB (1990) Asymptotic of solution to the equation of radiative transfer in a disk of large radius with a source near the symmetry axis, vol 97. Preprint of the Inst. Prikl. Matem., Akad. Nauk SSSR, Moscow
Peierls R (1939) Critical conditions in neutron multiplication. Proc Camb Phil Soc 35(4):610–615
Peraiah A (1973) Spectral line formation in extended atmospheres. Mon NoticRoy Astron Soc 162(4):321–327
Peraiah A, Grant I (1973) Numerical solution of the radiative transfer equation in spherical shells. J Inst Math Appl 12(1):75–90
Peraiah A (1975) Numerical solution of radiative transfer equation in extended spherical atmospheres with Rayleigh phase function. Astron Astrophys 40(1–2):75–80
Peraiah A (1978) Line formation in spherical media with partial frequency redistribution. I. Solution of the line transfer. Astrophys Space Sci 58(1):189–205
Peraiah A, Varghese BA (1984) Radiative transfer equation in spherically symmetric non-scattering media. Astrophys Space Sci 107(1):177–190
Peraiah A, Varghese BA (1985a) Solution of radiative transfer equation with spherical symmetry in partially scattering medium. Astrophys Space Sci 108(1):67–80
Peraiah A, Varghese BA (1985b) Radiative transfer equation in spherical symmetry. Astrophys J 290(2):411–423
Pikidjian HV (1978) On the Green's function of a non-coherent, anisotropically scattering plane-parallel slab. Astrofizika 14(1):169–190
Poon PTY, Ueno S (1974) Sobolev's Φ-function of radiative transfer in planar and spherical scattering media. Astrophys Space Sci 28(1):233–244
Pustylnick IB (1964) On the theory of a grey extended atmosphere. Publ Tartu Astrophys Obs 34:381–395
Pustylnick IB (1965) On the problem of continuous spectrum for the stars with extended atmospheres. Publ Tartu Astrophys Obs 35:138–152
Rogovtsov NN (1980) Reconstruction of the internal radiation field distribution from its characteristics at the scattering medium boundaries. Isvestiya, Atmos Ocean Phys 16:160–165
Rogovtsov NN (1981a) The transfer theory and the general invariance principle. Dokl Akad Nauk BSSR 25:420–423

Rogovtsov NN (1981b) Partial invariant determination and the relation between the radiation fields within a scattering volume and at the boundaries. J Appl Spectrosc 34:241–246
Rogovtsov NN (1983) Invariance relations and radiative transfer problems for complex configuration media. Dokl Akad Nauk BSSR 27:34–37
Rogovtsov NN (1985) On calculation of the radiation field characteristics in complex form scattering objects on the basis of the general invariance relation. Izd AkadNauk SSSR Fiz Atmos Okeana 21:1111–1112
Rogovtsov NN, Samson AM (1985) Asymptotical expressions for the internal radiation fields in plane-parallel media and applications to the calculation of mean emission durations of a layer and spherical shell. Astrophysics 23:468–476
Rogovtsov NN (1986a) Radiative transfer in an optically thick sphere. Zh Prikl Spectrosc 43:659–663
Rogovtsov NN (1986b) Asymptotical formulae for radiation fields in spherically and cylindrically symmetrical optically thick scattering media. Dokl Akad Nauk BSSR 30:901–904
Rogovtsov NN (1990) On asymptotical formulas for the radiative transfer equation Green's function for scattering media of a cylindrical form. Dokl Akad Nauk BSSR 34:1081–1084
Rogovtsov NN (1992a) On asymptotical formulas of the radiative transfer equation Green's function for scattering media of a layer form or a rotating cylinder form. Dokl Akad Nauk Belarus 36:26–29
Rogovtsov NN (1992b) On the derivation of the radiative transfer equation surface Green's function for the case of cylindrical form medium. Dokl Akad Nauk Belarus 36:561–598
Rogovtsov NN (1999) Properties and principles of invarince. Applications to solving mathematical physics problem, part 1, Education Ministry of RB, Minsk
Rogovtsov NN, Borovik FN (2009) The characteristic equation of radiative transfer theory. In: Kokhanovsky A (ed) Light scattering reviews, vol 4. Springer-Praxis, Chichester, pp 347–429
Rogovtsov NN (2010) General invariance relations reduction method and its applications to solutions of radiative transfer problems for turbid media of various configurations. In: Kokhanovsky A (ed) Light scattering reviews, vol 5. Springer-Praxis, Chichester, pp 249–327
Rogovtsov NN, Borovik FN, Kokhanovsky AA (2010) An effective algorithm of calculation the reflection function for the case of an arbitrary phase function. In: Abstracts NATO advanced study institute of special detection technique (polarimetry) and remote sensing, 12–25 Sept, Kyiv (Ukraine), p 99
Rogovtsov NN (2015a) Constructive theory of scalar characteristic equations of the theory of radiation transport: I. Basic assertions of the theory and conditions for the applicability of the truncation method. Diff Equ 51(2):268–281
Rogovtsov NN (2015b) Constructive theory of scalar characteristic equations of the theory of radiation transport: II. Algorithms for finding solutions and then analytic representations. Diff Equ 51(5):661–673
Rybicki GB (1970) A note on the computation of diffuse reflection function for spherical shells. J Comp Phys 6(1):131–135
Rybicki GB, Hunmer DG (1975) A note on the "peaking effect" in spherical-geometry transfer problems. Mon Notic Roy Astron Soc 170(1):423–427
Sabashvili SA (1973) Multiple scattering of resonance radiation in a plane layer and in a sphere. Astrofizika 9(2):273–292
Sahni DC (1975) Neutron transport problems in a spherical shell. J Math Phys 16(11):2260–2270
Sanford MT (1973) Radiative transfer in spherically symmetric circumstellar dust envelope. Astrophys J 183(2, pt.1):555–564
Schmid-Burgk J (1973) Integral kernels and exact solutions to some radiative problems in spherical geometries. Astrophys J 181(3, pt.1):865–874
Schmid-Burgk J (1975) Radiative transfer through spherically-symmetric atmospheres and shells. Astron Astrophys 40(3):249–255
Schmid-Burgk J, Scholz M (1975) Extended static stellar atmospheres. II. Location in the HR-diagram and some properties of extended atmosphere stars. Astron Astrophys 41(1):41–45

Schwarzschild K (1914) ÜberDiffusion and Absorption in der Sonnenatmosphäre. Sitzber. Akad. Wissenseh., Berlin, pp 1183–1202
Sen KK, Wilson SJ (1993) Generalized Eddington approximation for radiative transfer problems in sphericallysymmetric moving media. Astrophys Space Sci 203(2):227–240
Sheaks OJ (1972) Particle transport in spherical medium with a central black cavity. J Math Phys 13(2):203–209
Shikhov SB, Troyanovski VB (1983) Theory of nuclear reactors, vol. 2. Gas kinetic theory. Energoatomizdat, Moscow
Shkurpelov AA, Ershov YI (1967) Solution of the one-speed equation of neutron transport in a spherically symmetric medium. Zh Vychisl Mat i Mat Fiz 7(6):258–264
Shkurpelov AA, Shikhov SB, Ershov YuI (1970) Solution of the single-velocity kinetic equation of neutron transport in spherical geometry with scatter indicatrix linearly dependent on the angle. USSR Comp Matem Matem Fiz 10(3):257–264
Siewert CE, Benoist P (1979) The S_N-method in neutron-transport theory. I. Theory and applications. Nucl Sci Eng 69(2):156–160
Siewert CE, Grandjean P (1979) Three-basic neutron-transport problems in spherical geometry. Nucl Sci Eng 70(1):96–110
Siewert CE, Maiorino JR (1979) A point source in a finite sphere. J Quant Spectrosc Radiat Transfer 22(5):435–439
Siewert GE, Thomas JR (1984) Neutron transport calculations in cylindrical geometry. Nucl Sci Eng 87(2):107–112
Siewert GE, Thomas JR (1985) Radiative transfer calculations in spheres and cylinders. J Quant Spectrosc Radiat Transfer 34(1):59–64
Silant'ev NA (2006) Polarization of radiation scattered in magnetized turbulent envelopes. Astron Astrophys 449(2):597–608
Simonneau E (1976) Radiative transfer in atmospheres with spherical symmetry. J Quant Spectrosc Radiat Transfer 16(9):741–753
Simonneau E (1978) Radiative transfer in atmospheres with spherical symmetry. II. The conservative Milne problem. J Quant Spectrosc Radiat Transfer 19(5):439–450
Simonneau E (1980) Radiative transfer in atmospheres with spherical symmetry. IV. The non-conservative problem. J Quant Spectrosc Radiat Transfer 23(1):73–81
Smith MG (1965) The isotropic scattering of radiation from a point source in a finite spherical atmosphere. Proc Camb Phil Soc 61(4):923–937
Smoktii OI (1986) Modelling of radiation fields in problems of space spectrophotometry. Izd. Nauka, Leningrad
Smoktii OI, Anikonov AS (2008) Light scattering in media of large optical thickness. Izd. Nauka, St. Petersburg
Sobolev VV (1943) The approximation solution to the problem of light scattering in a medium with an arbitrary phase function. Astron Zh 20(5–6):14–22
Sobolev VV (1952) On the theory of the non-stationary radiation field. Astron Zh 29(4):406–414, 29(5):517–525
Sobolev VV (1958) Diffusion of radiation in a plane layer. Soviet Physics Doklady 3:541–544
Sobolev VV (1963) A treatise on radiative transfer. Van Nostrand, Prinston
Sobolev VV (1965) Diffusion on radiation in a nebula of large optical thickness. Kinematics and dynamics of stellar systems and physics of interstellar medium. Izd. Nauka, Alma-Ata, pp 285–291
Sobolev VV (1972) Light scattering in a homogeneous sphere. Astrophysics 8(2):117–127
Sobolev VV (1975) Light scattering in planetary atmospheres. Pergamon Press, Oxford
Sobolev VV (1983) Scattering of light in a spherical shell and a sphere. Soviet Phys Dokl 28:908–911
Sobolev VV (1984) Integral relations and asymptotic expressions in the theory of radiative transfer. Astrophysics 20(1):79–85
Takenti M (1979) On Unno-Kondo's generalized Eddington approximation in extended atmospheres. Publ Astron Soc Jpn 31(1):199–202

Taylor PA (1927) The equilibrium of the calcium chromosphere. Mon Notic Roy Astron Soc 87(8):605–616
Tsujita J (1967) Diffuse reflection and transmission by an inhomogeneous infinite cylindrical region. Publ Astron Soc Jpn 19(3):468–475
Tsujita J (1969) The invariant imbedding equation for the dissipation function of an infinite inhomogeneous cylindrical shell. Publ Astron Soc Jpn 21(1):15–20
Ueno S, Kagiwada H, Kalaba R (1971) Radiative transfer in spherical shell atmosphere with radial symmetry. J Math Phys 12(7):1279–1286
Ueno S (1974) On the resolvent of Milne's integral equation for a spherical isotropically-scattering medium. J Quant Spectrosc Radiat Transfer 14(4):245–249
Unno W, Kondo M (1976) The Eddington approximation generalized for radiative transfer in spherically symmetric systems. I. Basic method. Publ Astron Soc Jpn 28(2):347–354
Unno W, Kondo M (1977) The Eddington approximation generalized for radiative transfer in spherically symmetric systems. II. Nongrey extended dust-shell models. Publ Astron Soc Jpn 29(4):693–710
Van de Hulst HC (1980) Multiple light scattering. Tables, Formulas, and applications, vols. 1, 2. Academic Press, New York
Van de Hulst HC (1987) Radiative transfer in a spherical dust cloud. I. Exact results for isotropic scattering. Astron Astrophys 173(1):115–123
Van de Hulst HC (1988) Radiative transfer in a spherical dust cloud. II. Asymptotic form of the reflection function for isotropic scattering. Astron Astrophys 207(1):182–193
Van de Hulst HC (1994) Radiative transfer by a homogeneous spherical cloud. Trudy Astron Obs St Petersburg Gos Univ 44:131–153
Viik T (1974a) Radiative transfer in homogeneous spherical shell atmospheres. Publ Tartu Astrophys Obs 42:3–21
Viik T (1974b) A modified regional averaging method. Publ Tartu Astrophys Obs 42:22–37
Viik T (1974c) A modified regional averaging method. II. Comparision with the P-3 approximation. Publ Tartu Astrophys Obs 42:281–292
Vladimirov VS (1963) Mathematical problems in the one-velocity theory of particle transport. Atomic Energy of Canada Limited, Chalk River
Vladimirov VS (1976) Distributions in mathematical physics. Izd. Nauka, Moscow
Weinberg A, Wigner E (1958) The physical theory of neutron chain reactors. University of Chicago Press, Chicago
Wick GC (1943) Über ebene Diffusionsprobleme. Zeitschr Phys 121(11–12):702–718
Wilson SJ, Sen KK (1964) Modified spherical harmonic method and spherical geometry. I. Diffusion problem. Ann Astrophys 27(6):654–659
Wilson SJ, Sen KK (1965a) Modified spherical harmonic method and spherical geometry. II. Transfer problem in spherically symmetric finite stellar atmosphere. Ann Astrophys 28(2):348–352
Wilson SJ, Sen KK (1965b) Modified spherical harmonic method and spherical geometry. III. Transfer problem in spherically symmetric finite shell of planetary nebulae. Ann Astrophys 28(5):855–859
Wilson SJ, Tung CT, Sen KK (1972) The Eddington factor for radiative transfer in spherical geometry. Mon Notic Roy Astron Soc 160(4):349–353
Wilson SJ, Sen KK (1973a) On the resolvent kernel for the transfer problem in an homogeneous sphere. J Quant Spectrosc Radiat Transfer 13(1):83–85
Wilson SJ, Sen KK (1973b) Probabilistic model for the resolvent kernel in diffusion problems in spherical-shell media. J Quant Spectrosc Radiat Transfer 13(3):255–266
Wilson SJ, Sen KK (1973c) Probabilistic model for the resolvent kernel in non-coherent scattering sphericalshell. J Quant Spectrosc Radiat Transfer 13(8):755–763
Wilson SJ, Wan FS, Sen KK (1980) On the moment method for radiative transfer in spherical dust shells with a core. Astrophys Space Sci 67(1):99–108
Wing GM (1962) An introduction to transport theory. Wiley, New York
Yanovitski EG (1997) Light scattering in inhomogeneous atmospheres. Springer, New York

Yengibarian NB (1972) The scattering of light in a sphere with arbitrary source distribution. Astrofizika 8(1):149–153
Yorke HW (1986) Numerical solution of the equation of radiative transfer in spherical geometry. Ann Astrophys 86(3):286–297
Yvon J (1957) La diffusion mactroscopique des neutrons: une methode d'approximation. J Nucl Energy 4(3):305–318

Part II
Single Light Scattering

The Debye Series and Its Use in Time-Domain Scattering

James A. Lock and Philip Laven

Abstract In electromagnetic scattering of an incident beam by a single particle possessing a reasonably high degree of symmetry, the Debye series decomposes the partial wave scattering and interior amplitudes into the contributions of a number of intuitive physical processes. We describe the Debye series for scattering by a sphere, a coated sphere, a multi-layer sphere, a tilted cylinder, and a prolate spheroid. We also comment on the meaning of the various Debye series terms, and briefly recount the method by which the formulas of ray scattering can be derived from them. We also consider time-domain scattering of a short pulse by a single particle and describe the way in which the time-domain scattering signature naturally separates out the various Debye series terms. Lastly, we show how time-domain scattering further separates a number of cooperating sub-processes present in individual Debye series terms.

1 Introduction

This monograph is concerned with the physical interpretation of scattering of an incident electromagnetic beam by a single particle of relatively high symmetry, such as a sphere, a circular cylinder, or a spheroid. If one is to successfully interpret the results of experiments that use light scattering as a nonintrusive technique for particle characterization, one must first understand the physical meaning of the many complicated features present in the scattering signature. As an example, for scattering of a plane wave of wavelength λ by a spherical particle of radius a in air, the scattered intensity has approximately $2\pi a/\lambda$ oscillations between forward scattering and back-scattering. Each oscillation has a specific meaning, and results from the action of one or more scattering mechanisms. However, the mathematical complexity of the equations of Lorenz–Mie scattering of an electromagnetic plane

J.A. Lock (✉)
Department of Physics, Cleveland State University, Cleveland, OH 44115, USA
e-mail: j.lock@csuohio.edu

P. Laven
9 Russells Crescent, Horley RH6 7DJ, UK
e-mail: philip@philiplaven.com

© Springer-Verlag Berlin Heidelberg 2016

A. Kokhanovsky (ed.), *Light Scattering Reviews, Volume 11*,
Springer Praxis Books, DOI 10.1007/978-3-662-49538-4_5

wave gives no clue as to what those mechanisms might be [see Sects. 9.2, 9.3 of van de Hulst (1981), Sect. 3.3 of Kerker (1969), and Chap. 4 of Bohren and Huffman (1983)]. It will be shown in this monograph that almost all of these physical mechanisms can be readily understood by re-expressing the Lorenz–Mie partial wave scattering amplitudes a_n and b_n as an infinite series, known as the Debye series. In the remainder of this section, we recount both the history and the current state-of-the-art of the Debye series.

When summarizing the development of the Debye series over the last century, one must differentiate between scalar waves and electromagnetic waves, the type of incident beam considered, and the geometry of the scattering particle. The Debye series was first developed for scattering of a normally incident scalar plane wave by an infinitely long circular cylinder by Debye in 1908 in order to provide a rigorous derivation, starting from an exact wave equation, of the formulas of geometrical ray scattering. The derivation consisted of two basic steps: (i) expanding a mathematical expression as a geometric series and (ii) making the small wavelength approximation to get to the ray limit. Debye applied the small wavelength approximation to the partial wave scattering amplitudes first, and then expanded the result as a geometric series. In 1946 van de Hulst [as recounted on p. 226 of van de Hulst (1981)] considered an electromagnetic plane wave incident on a sphere, and again applied the small wavelength approximation before expanding the geometric series. Thus, the "Debye series" they derived is only an approximate decomposition of the partial wave scattering amplitudes, rather than being exact.

In 1937, van der Pol and Bremmer took the complementary point of view and expanded the partial wave scattering amplitudes for a scalar plane wave incident on a sphere as a geometric series first, before applying the small wavelength approximation to each of the terms individually, thus obtaining the exact Debye series familiar to researchers today. In 1954, Franz obtained an analogous exact Debye series for scattering of a scalar plane wave by an infinitely long circular cylinder at normal incidence. During the 1960s Rubinow (1961), Chen (1964), and Nussenzveig (1969a) applied the Watson transform to each of the exact Debye series terms individually for scattering of a scalar plane wave by a sphere in order to uncover a number of subtle and remarkably interesting wave corrections to the dominant ray theory results. This same process was repeated by Khare in 1976 and Nussenzveig in 1979 for an electromagnetic plane wave incident on a sphere. The form of the Debye series these authors used was well suited for the analytical calculation of the various wave scattering corrections. But it was not optimally suited for numerical computations.

Beginning in the late 1980s, the exact Debye series was derived for electromagnetic plane wave incidence for a number of different scattering geometries in a convenient form that is easily programmable for numerical computations (Lock 1988; Hovenac and Lock 1992). These include the Debye series of the interior partial wave amplitudes (Hovenac and Lock 1992; Lock and Hovenac 1991), a sphere with an embedded electric dipole source (Lock 2001), coated and multi-layer spheres (Lock et al. 1994; Lock 2005, 2008; Li et al. 2006a; Lock and Laven 2012), diagonal plane wave incidence on an infinitely long circular cylinder (Lock and Adler 1997), and normal (Li et al. 2006b) and diagonal (Li et al. 2009) plane wave incidence on a multi-layer

cylinder. For an incident electromagnetic Gaussian beam, the Debye series has been applied to scattering by a sphere (Lock 1993; Gouesbet 2003), a multi-layer sphere (Li et al. 2007), and for normal incidence on a multi-layer circular cylinder (Wu and Li 2008). The exact Debye series has also been derived for other beam and particle types, including an electromagnetic Neumann beam incident on a sphere (Lock 2011), and an electromagnetic plane wave incident on a spheroid (Xu et al. 2010a), a general non-spherical particle (Xu et al. 2010b), a coated non-spherical particle (Xu and Lock 2010), and a highly absorbing sphere (Shen and Wang 2010).

In each of these situations, the Debye series is not a short-wavelength approximation to the partial wave scattering amplitudes. Rather, it is an alternate form of the exact amplitudes themselves, which is valid for particles of all sizes. In the language of frequency-domain multiple scattering, the Lorenz–Mie formulas can be thought of as the compact form of the partial wave scattering amplitudes, whereas the Debye series is the expanded form (Twersky 1964). A separate question is the following: although the Debye series is always exact, is it also always useful for deriving physical intuition concerning the scattering processes? The answer is that the ray-like physical mechanisms suggested by the Debye series are at least qualitatively valid when $2\pi a/\lambda > 50$ (Glantschnig and Chen 1981; Ungut et al. 1981).

The body of this monograph is organized as follows. In Sect. 2, we summarize the equations of Lorenz–Mie scattering for an electromagnetic plane wave incident on a spherical particle. The derivation of the Debye series for the partial wave scattering and interior amplitudes for a sphere proceeds in two steps, which are described in Sects. 3 and 4. These results are then used in Sect. 5 to rigorously derive the ray theory approximation to scattering by a sphere in the $\lambda \ll a$ limit. The Debye series for scattering by a sphere is expressed in terms of four fundamental partial wave amplitudes. But only two of them appear in the compact form of the partial wave scattering amplitudes. Section 6 describes another incident beam type that makes use of the other two fundamental amplitudes for the compact form. The Debye series for an electromagnetic plane wave incident on coated sphere and a multi-layer sphere is described in Sect. 7. We find that the results are formally identical to those of a homogeneous sphere when they are expressed in terms of multiple-scattering amplitudes rather than in terms of single-scattering amplitudes. This gives rise to an expanded and compact form of the Debye series itself. Scattering of a diagonally incident electromagnetic plane wave by an infinitely long circular cylinder is treated in Sect. 8. As opposed to scattering by a sphere, the Debye series here contains both polarization-preserving and polarization-changing components. The Debye series for scattering of a scalar plane wave and an electromagnetic plane wave by a prolate spheroid is described in Sect. 9. The new feature for this geometry is that the terms of the Debye series exhibit coupling of each partial wave of the incident beam to all the partial waves of the scattered and interior waves. In Sect. 10, we describe time-domain scattering of a temporally short electromagnetic plane wave pulse by a sphere, and show how it separates the many different scattered signals that occur at the same scattering angle in frequency-domain scattering as in Sects. 2–9. In Sect. 11, we find that each term of the Debye series carries a different and easily identifiable temporal signature in

time-domain scattering. In Sect. 12, we examine the time-domain signature of a number of subtle and delicate wave scattering corrections to the dominant ray theory results for $a \gg \lambda$ that are implicit in the various terms of the Debye series. Finally, in Sect. 13 we give a few concluding thoughts concerning the significance of the Debye series. In each of these sections, with the exception of Sects. 4 and 5, the results are only quoted rather than derived in detail. The derivation in Sect. 4 is important for extending the Debye series from a spherical particle to more general situations. A straightforward derivation of ray theory from the Debye series as is given in Sect. 5, to the best of our knowledge, has not appeared in the recent literature, and is crucial for the understanding of the results of Sect. 12.

2 Scattering by a Homogeneous Sphere in Lorenz–Mie Theory

The standard Lorenz–Mie scattering problem considers an electromagnetic plane wave of vacuum wavelength λ, wave number $k = 2\pi/\lambda$, electric field amplitude E_0 in an external medium of refractive index m_2, propagating in the $+z$ direction, and linearly polarized in the x direction. It is scattered by a homogeneous dielectric spherical particle of radius a, refractive index m_1, whose center is located at the origin of coordinates. The external medium will hereafter be denoted as region 2 and the particle interior will be region 1, except in Sect. 7 which is concerned with scattering by a coated sphere and a multi-layer sphere for which there are a number of concentric regions. The time dependence $\exp(-i\omega t)$ of the incident, scattered, and interior fields will be left implicit, where $\omega = ck$ and c is the speed of light in vacuum.

The transverse electric polarized (TE) and transverse magnetic polarized (TM) scalar radiation potential of the incident, scattered, and interior electromagnetic waves is denoted by $\psi^X(m_j kr, \theta, \varphi)$ in spherical coordinates (r, θ, φ) with respect to the origin, where m_j is the refractive index of region j and X = TE, TM. It satisfies the scalar wave equation, also known as the scalar Helmholtz equation,

$$\nabla^2 \psi^X + (m_j k)^2 \psi^X = 0. \tag{2.1}$$

The solution of Eq. (2.1) is written as a sum over the partial wave number $1 \le n < \infty$ and integer azimuthal mode number m with $-n \le m \le n$ of the function $\psi^X_{n,m}$ multiplied by an arbitrary complex constant which gives the magnitude and phase of the partial wave contribution. The function $\psi^X{}_{n,m}$ contains a spherical Bessel, Neumann, or Hankel function of $m_j kr$ multiplied by an associated Legendre function $P_n^m[\cos(\theta)]$, and the complex exponential $\exp(im\varphi)$. The TE-polarized electric and magnetic fields associated with $\psi^{\mathrm{TE}}_{n,m}$ are [see p. 120 of van de Hulst (1981) and Sect. 7.11 of Stratton et al. (1941)]

$$\mathbf{E}^{\mathbf{TE}} = -\mathbf{r} \times \nabla \psi^{\mathrm{TE}}{}_{n,m} \tag{2.2a}$$

$$c\mathbf{B}^{\mathbf{TE}} = \left(-\mathrm{i}/m_j k\right) \nabla \times \mathbf{E}^{\mathbf{TE}}, \tag{2.2b}$$

and the TM-polarized electric and magnetic fields are

$$\left(c/m_j\right) \mathbf{B}^{\mathbf{TM}} = \mathbf{r} \times \nabla \psi^{\mathrm{TM}}_{n,m} \tag{2.3a}$$

$$\mathbf{E}^{\mathbf{TM}} = \left[\mathrm{i}c/\left(m_j^2 k\right)\right] \nabla \times \mathbf{B}^{\mathbf{TM}}. \tag{2.3b}$$

As an alternative notation, Eqs. (2.2a, 2.3a) are frequently associated with the vector spherical wave function $\mathbf{M_{m,n}}(k\mathbf{r})$, and Eqs. (2.2b, 2.3b) are associated with $\mathbf{N_{m,n}}(k\mathbf{r})$. As an aside, we note that the usual convention for the vector spherical wave functions uses the azimuthal mode as the first subscript and the partial wave number as the second subscript.

The vector derivatives in Eqs. (2.2a, 2.2b, 2.3a, 2.3b), when expressed in spherical coordinates, make it convenient to use scalar radiation potentials in the Riccati–Bessel form

$$\Psi^X_{n,m}\left(m_j kr, \theta, \varphi\right) = m_j kr\, \psi^X_{n,m}\left(m_j kr, \theta, \varphi\right). \tag{2.4}$$

For an x-polarized electromagnetic plane wave of field strength E_0 traveling in the positive z direction and incident on a spherical particle, only the azimuthal modes $m = \pm 1$ contribute. In the notation of Eq. (2.4), the scalar radiation potential of the TE portion of the incident plane wave is taken to be

$$\Psi^{\mathrm{inc,TE}}(r, \theta, \varphi) = m_2 E_0 \sum_{n=1}^{\infty} C_n \psi_n(m_2 kr) P_n{}^1[\cos(\theta)] \sin(\varphi), \tag{2.5}$$

where

$$C_n = \mathrm{i}^n (2n+1)/[n(n+1)], \tag{2.6}$$

and $\psi_n(m_2 kr)$ is a Riccati–Bessel function [not to be confused with the scalar radiation potential $\psi^X_{n,m}\left(m_j kr, \theta, \varphi\right)$ of Eq. (2.4)]

$$\psi_n(z) = z j_n(z), \tag{2.7}$$

with j_n being a spherical Bessel function. The TE portion of the scattered and interior scalar radiation potential is taken to be

$$\Psi^{\mathrm{scatt,TE}}(r, \theta, \varphi) = -m_2 E_0 \sum_{n=1}^{\infty} C_n b_n \zeta_n^{(1)}(m_2 kr) P_n^1[\cos(\theta)] \sin(\varphi) \tag{2.8a}$$

$$\Psi^{\text{inside,TE}}(r,\theta,\varphi) = m_1 E_0 \sum_{n=1}^{\infty} C_n d_n \psi_n(m_1 kr) P_n^1[\cos(\theta)] \sin(\varphi), \tag{2.8b}$$

where b_n and d_n are the TE partial wave scattering and interior amplitudes, and $\zeta_n^{(1)}$ is a Riccati–Hankel function of the first kind

$$\zeta_n^{(1)}(z) = z h_n^{(1)}(z), \tag{2.9}$$

with $h_n^{(1)}$ being a spherical Hankel function of the first kind which describes a radially outgoing traveling wave when the time dependence $\exp(-\mathrm{i}\omega t)$ is appended. The expressions for the TM scalar radiation potentials are similar, but contain the TM partial wave scattering and interior amplitudes a_n and c_n, and $\cos(\varphi)$ rather than $\sin(\varphi)$. The boundary conditions on the tangential and normal components of the total electric and magnetic field in each region that must be satisfied at the surface of the sphere are equivalent to the continuity of $\Psi^{\text{TE}}/(m_j ka)$ and $\Psi^{\text{TE}'}$ for the TE waves, and the continuity of Ψ^{TM} and $\Psi^{\text{TM}'}/(m_j ka)$ for the TM waves, where the derivative of Ψ with respect to its radial argument is denoted as Ψ'.

The resulting Lorenz–Mie scattered fields in the far-zone $r \gg a$ are found to be (van de Hulst 1981; Kerker 1969; Bohren and Huffman 1983)

$$\mathbf{E}^{\text{scatt}}(r,\theta,\varphi) = [\mathrm{i}E_0 \exp(\mathrm{i}m_2 kr)/(m_2 kr)]\left[S_2(\theta)\cos(\varphi)\mathbf{u}_{\boldsymbol{\theta}} - S_1(\theta)\sin(\varphi)\mathbf{u}_{\boldsymbol{\varphi}}\right] \tag{2.10a}$$

$$(c/m_2)\mathbf{B}^{\text{scatt}}(r,\theta,\varphi) = [\mathrm{i}E_0 \exp(\mathrm{i}m_2 kr)/(m_2 kr)]\left[S_1(\theta)\sin(\varphi)\mathbf{u}_{\boldsymbol{\theta}} + S_2(\theta)\cos(\varphi)\mathbf{u}_{\boldsymbol{\varphi}}\right]. \tag{2.10b}$$

The scattering amplitudes $S_1(\theta)$ and $S_2(\theta)$ are the partial wave sums

$$S_1(\theta) = \sum_{n=1}^{n_{\max}} \{(2n+1)/[n(n+1)]\}\,[a_n \pi_n(\theta) + b_n \tau_n(\theta)] \tag{2.11a}$$

$$S_2(\theta) = \sum_{n=1}^{n_{\max}} \{(2n+1)/[n(n+1)]\}\,[a_n \tau_n(\theta) + b_n \pi_n(\theta)], \tag{2.11b}$$

where the Lorenz–Mie angular functions are:

$$\pi_n(\theta) = [1/\sin(\theta)]P_n^1[\cos(\theta)] \tag{2.12a}$$

$$\tau_n(\theta) = (\mathrm{d}/\mathrm{d}\theta)P_n^1[\cos(\theta)]. \tag{2.12b}$$

The partial wave sum in Eqs. (2.11a, 2.11b) has been found to be numerically convergent to single precision accuracy when (Wiscombe 1980)

$$n_{\max} = x + 4.2x^{1/3} + 1, \tag{2.13}$$

where the size parameter of the sphere is defined as

$$x = 2\pi m_2 a/\lambda. \tag{2.14}$$

Since we will show in Sect. 5 that $\tau_n(\theta)$ is usually much larger than $\pi_n(\theta)$ when $a \gg \lambda$, $S_1(\theta)$ will hereafter be called the TE scattering amplitude and $S_2(\theta)$ will be called the TM scattering amplitude. Equations (2.10a, 2.10b) then indicate that only TE scattering occurs in the yz plane and only TM scattering occurs in the xz plane.

The partial wave scattering and interior amplitudes and the terms in their Debye series expansion are conveniently written in terms of four fundamental partial wave amplitudes (Lock 1988)

$$N_n = \alpha\,\psi_n(m_2ka)\,\psi'_n(m_1ka) - \beta\,\psi'_n(m_2ka)\,\psi_n(m_1ka) \tag{2.15a}$$

$$D_n = \alpha\,\chi_n(m_2ka)\,\psi'_n(m_1ka) - \beta\,\chi'_n(m_2ka)\,\psi_n(m_1ka) \tag{2.15b}$$

$$P_n = \alpha\,\psi_n(m_2ka)\,\chi'_n(m_1ka) - \beta\,\psi'_n(m_2ka)\,\chi_n(m_1ka) \tag{2.15c}$$

$$Q_n = \alpha\,\chi_n(m_2ka)\,\chi'_n(m_1ka) - \beta\,\chi'_n(m_2ka)\,\chi_n(m_1ka), \tag{2.15d}$$

where χ_n is a Riccati–Neumann function

$$\chi_n(z) = z\,n_n(z), \tag{2.16}$$

with n_n being a spherical Neumann function. In addition, $\alpha = m_1$ and $\beta = m_2$ for the TE polarization and $\alpha = m_2$ and $\beta = m_1$ for the TM polarization. The four fundamental partial wave amplitudes satisfy the Wronskian relation [see p. 631 of Arfken (1985)]

$$N_nQ_n - D_nP_n = m_1m_2. \tag{2.17}$$

The Lorenz–Mie partial wave scattering amplitudes may be written in terms of N_n and D_n as

$$a_n, b_n = N_n/(N_n + \mathrm{i}\,D_n) = U_n^{-1}N_n, \tag{2.18}$$

where

$$U_n = N_n + \mathrm{i}\,D_n. \tag{2.19}$$

In Eq. (2.18) the reciprocal of U_n is written as U_n^{-1} in order to be consistent in notation with the matrix version of the equation that will be obtained in Sect. 9. Lastly, the fields inside the spherical particle are

$$\begin{aligned}\mathbf{E}^{\mathbf{inside}}(r,\theta,\varphi) = &-\mathrm{i}(m_1/m_2)E_0\sum_{n=1}^{n_{\max}}\mathrm{i}^n\,(2n+1)c_n\,[j_n(m_1kr)/(m_1kr)]P_n^1[\cos(\theta)]\cos(\varphi)\mathbf{u_r}\\ &+(m_1/m_2)E_0\sum_{n=1}^{n_{\max}}\mathrm{i}^n\,\{(2n+1)/[n(n+1)]\}\,\{d_nj_n(m_1kr)\,\pi_n(\theta)\\ &-\mathrm{i}c_n\left[\psi_n'(m_1kr)/(m_1kr)\right]\,\tau_n(\theta)\}\cos(\varphi)\mathbf{u_\theta}\\ &-(m_1/m_2)E_0\sum_{n=1}^{n_{\max}}\mathrm{i}^n\,\{(2n+1)/[n(n+1)]\}\{d_nj_n(m_1kr)\tau_n(\theta)\\ &-\mathrm{i}c_n\left[\psi_n'(m_1kr)/(m_1kr)\right]\,\pi_n(\theta)\}\,\sin(\varphi)\mathbf{u}_\varphi\end{aligned} \tag{2.20a}$$

$$\begin{aligned}(c/m_1)\mathbf{B}^{\mathbf{inside}}(r,\theta,\varphi) = &-\mathrm{i}(m_1/m_2)E_0\sum_{n=1}^{n_{\max}}\mathrm{i}^n(2n+1)d_n[j_n(m_1kr)/(m_1kr)]P_n^1[\cos(\theta)]\sin(\varphi)\mathbf{u_r}\\ &+(m_1/m_2)E_0\sum_{n=1}^{n_{\max}}\mathrm{i}^n\{(2n+1)/[n(n+1)]\}\{c_nj_n(m_1kr)\pi_n(\theta)\\ &-\mathrm{i}d_n\left[\psi_n'(m_1kr)/(m_1kr)\right]\tau_n(\theta)\}\sin(\varphi)\mathbf{u_\theta}\\ &+(m_1/m_2)E_0\sum_{n=1}^{n_{\max}}\mathrm{i}^n\{(2n+1)/[n(n+1)]\}\{c_nj_n(m_1kr)\tau_n(\theta)\\ &-\mathrm{i}d_n\left[\psi_n'(m_1kr)/(m_1kr)\right]\pi_n(\theta)\}\cos(\varphi)\mathbf{u}_\varphi,\end{aligned} \tag{2.20b}$$

where the interior partial wave scattering amplitudes may be written as

$$c_n, d_n = -\mathrm{i}m_2/(N_n + \mathrm{i}\,D_n) = -\mathrm{i}m_2\,U_n^{-1}. \tag{2.21}$$

As an example illustrating the complicated structure of the TE and TM scattered intensity, Fig. 1 shows the TE intensity $|S_1(\theta)|^2$ and the TM intensity $|S_2(\theta)|^2$ for λ = 0.65 μm, a = 10 μm, m_1 = 1.33257, and m_2 = 1, corresponding to a spherical water droplet of size parameter x = 96.66 in air illuminated by monochromatic visible light. This structure will be explained in terms of the Debye series in Sect. 4.

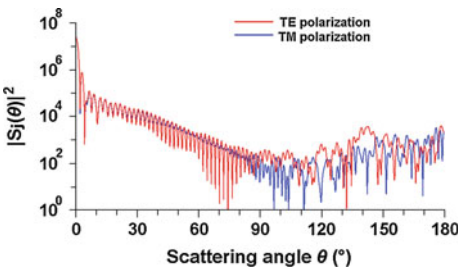


Fig. 1 TE and TM intensity as a function of angle for scattering of an x-polarized plane wave by a sphere with λ = 0.65 μm, a = 10 μm, m_1 = 1.33257, and m_2 = 1

3 Partial Wave Fresnel Coefficients for Scattering by a Homogeneous Sphere

The first step in the derivation of the Debye series is to determine the magnitude and phase of the transmitted and reflected waves when a single radially incoming or outgoing partial wave is incident on the sphere surface from either region 2 or 1, respectively. Consider first the scattering problem where the beam is a single radially incoming spherical multipole wave of unit amplitude and partial wave number n in region 2 incident on the sphere surface from the outside. The scalar radiation potential for the TE version of the incoming wave is

$$\Psi_n^{\mathrm{inc,TE}}(r,\theta,\varphi) = C_n \zeta_n^{(2)}(m_2 kr) P_n^1[\cos(\theta)]\sin(\varphi) \tag{3.1}$$

where $\zeta_n^{(2)}$ is a Riccati–Hankel function of the second kind

$$\zeta_n^{(2)}(z) = z\, h_n^{(2)}(z), \tag{3.2}$$

with $h_n^{(2)}$ being a spherical Hankel function of the second kind which describes a radially incoming traveling wave when the time dependence $\exp(-\mathrm{i}\omega t)$ is appended. Part of the incident wave, having the complex amplitude T_n^{21}, is transmitted into the sphere as a radially incoming partial wave in region 1 (reading the superscripts from left to right, i.e., incident from region 2 into region 1), and part of it, having the complex amplitude R_n^{212}, is reflected back into region 2 as a radially outgoing partial wave (again reading the superscripts from left to right, i.e., incident from 2, and reflecting at 1 back into 2),

$$\Psi_n^{\mathrm{ref,TE}}(r,\theta,\varphi) = C_n R_n^{212} \zeta_n^{(1)}(m_2 kr) P_n^1[\cos(\theta)]\sin(\varphi) \tag{3.3a}$$

$$\Psi_n^{\mathrm{trans,TE}}(r,\theta,\varphi) = C_n T_n^{21} \zeta_n^{(2)}(m_1 kr) P_n^1[\cos(\theta)]\sin(\varphi). \tag{3.3b}$$

For either polarization, matching the boundary conditions of the tangential and normal components of the electric and magnetic fields at the sphere surface gives

$$T_n^{21} = -2\mathrm{i}m_1/[(N_n+Q_n)+\mathrm{i}(D_n-P_n)] = -2\mathrm{i}m_1 M_n^{-1} \tag{3.4a}$$

$$R_n^{212} = [(-N_n+Q_n)+\mathrm{i}(D_n+P_n)]/[(N_n+Q_n)+\mathrm{i}(D_n-P_n)] = M_n^{-1}J_n, \tag{3.4b}$$

where

$$\begin{aligned} M_n &= (N_n+Q_n)+\mathrm{i}(D_n-P_n) \\ &= \alpha\,\zeta_n^{(1)}(m_2 ka)\zeta_n^{(2)\prime}(m_1 ka) - \beta\,\zeta_n^{(1)\prime}(m_2 ka)\zeta_n^{(2)}(m_1 ka) \end{aligned} \tag{3.5a}$$

$$J_n = (-N_n + Q_n) + \mathrm{i}(D_n + P_n)$$
$$= -\left[\alpha\, \zeta_n^{(2)}(m_2 ka)\zeta_n^{(2)\prime}(m_1 ka) - \beta\, \zeta_n^{(2)\prime}(m_2 ka)\zeta_n^{(2)}(m_1 ka)\right]. \tag{3.5b}$$

The derivation of the TM case is exactly the same, except that cos(φ) replaces sin(φ) in Eqs. (3.1, 3.3a, 3.3b).

We now apply this method to another scattering problem where the beam is a radially outgoing spherical multipole wave of unit amplitude and partial wave number n in region 1 incident on the sphere surface from the inside. Part of the incident wave, having the complex amplitude T_n^{12}, is transmitted out of the sphere as a radially outgoing partial wave in region 2 (reading the superscripts from left to right, i.e., incident from 1 into 2), and part of it, having the complex amplitude R_n^{121}, is reflected back into region 1 as a radially incoming partial wave (again reading the superscripts from left to right, i.e., incident from 1, and reflecting at 2 back into 1). The scalar radiation potential of the TE version for the incident, reflected, and transmitted waves is

$$\Psi_n^{\mathrm{inc,TE}}(r,\theta,\varphi) = C_n \zeta_n^{(1)}(m_1 kr) P_n^1[\cos(\theta)]\sin(\varphi) \tag{3.6a}$$

$$\Psi_n^{\mathrm{ref,TE}}(r,\theta,\varphi) = C_n R_n^{121} \zeta_n^{(2)}(m_1 kr) P_n^1[\cos(\theta)]\sin(\varphi) \tag{3.6b}$$

$$\Psi_n^{\mathrm{trans,TE}}(r,\theta,\varphi) = C_n T_n^{12} \zeta_n^{(1)}(m_2 kr) P_n^1[\cos(\theta)]\sin(\varphi). \tag{3.6c}$$

For either polarization, matching the boundary conditions of the tangential and normal components of electric and magnetic fields at the sphere surface gives

$$T_n^{12} = -2\mathrm{i}m_2/[(N_n + Q_n) + \mathrm{i}(D_n - P_n)] = -2\mathrm{i}m_2 M_n^{-1} = (m_2/m_1)T_n^{21} \tag{3.7a}$$

$$R_n^{121} = [(-N_n + Q_n) - \mathrm{i}(D_n + P_n)]/[(N_n + Q_n) + \mathrm{i}(D_n - P_n)] = M_n^{-1} L_n, \tag{3.7b}$$

where

$$L_n = (-N_n + Q_n) - \mathrm{i}(D_n + P_n)$$
$$= -\left[\alpha\, \zeta_n^{(1)}(m_2 ka)\zeta_n^{(1)\prime}(m_1 ka) - \beta\, \zeta_n^{(1)\prime}(m_2 ka)\zeta_n^{(1)}(m_1 ka)\right]. \tag{3.8}$$

Again the derivation of the TM case is exactly the same, except that cos(φ) replaces sin(φ) in Eqs. (3.6a–3.6c). These fractions of the incident amplitude can be thought of as the spherical multipole wave version of the Fresnel transmission and reflection coefficients for a plane wave incident on a flat interface [see Sect. 4.3.2 of Hecht (1987)]. This association will be developed further in Sect. 5. The partial wave reflection amplitudes have the same magnitude but differing phases (Lock 1988)

$$\left|R_n^{212}\right| = \left|R_n^{121}\right|, \tag{3.9a}$$

and the reflection and transmission amplitudes together satisfy the energy conservation relation

$$\left|R_n^{212}\right|\left|R_n^{121}\right| + \left|T_n^{21}\right|\left|T_n^{12}\right| = 1. \tag{3.9b}$$

4 Derivation of the Debye Series for Scattering by a Homogeneous Sphere

A Debye series expansion of the partial wave scattering and interior amplitudes can be developed whenever the scattering particle has a sufficiently high degree of symmetry so that a partial wave expansion can be used to calculate the scattered fields. As a side note, a Debye series expansion can still be carried out for arbitrary-shaped particles using a partial wave expansion in the context of T-matrix methods (Xu et al. 2010b; Xu and Lock 2010). The derivation in this section is carried out in two steps. First, all the partial wave transmission and reflection coefficients must be determined. This was accomplished in Sect. 3 for the case of scattering by a homogeneous sphere. The results are Eqs. (3.4a, 3.4b, 3.7a, 3.7b). For the second step, one must show that a_n, b_n, c_n, and d_n of Eqs. (2.18, 2.21) can be written solely in terms of these transmission and reflection coefficients, with nothing missing and nothing left over. For relatively simple geometries, such as a sphere, this can be accomplished by direct substitution (Lock 1988). But for more complicated geometries the polarization state of the light can either be maintained or change at the particle surface, as occurs for a plane wave diagonally incident on a circular cylinder (Lock and Adler 1997). A partial wave of the incident fields can also be coupled to all the partial waves of the scattered and interior fields, as occurs for scattering by a spheroid (Xu et al. 2010a). Thus, a more systematic version of the second step in the derivation is required in order to be applicable to these more complicated possibilities. The approach outlined here (Xu et al. 2010a) has the advantage that it can be straightforwardly generalized to the more complicated possibilities mentioned above.

In generalized Lorenz–Mie theory (GLMT) for scattering of an arbitrary-shaped beam, the scalar radiation potential for the TE version of the arbitrary incident beam is

$$\Psi^{\mathrm{inc,TE}}(r,\theta,\varphi) = m_2 \sum_{n=1}^{\infty} \sum_{m=-n}^{n} C_n G_{n,m}^{\mathrm{TE}} \psi_n(m_2 kr) P_n^m[\cos(\theta)] \exp(\mathrm{i}m\varphi), \tag{4.1}$$

where $G_{n,m}^{\mathrm{TE}}$ is the TE shape coefficient of the incident beam. The normalization coefficient C_n of Eq. (2.6) insures that scalar radiation potential of an x-polarized plane wave has the beam shape coefficients $G_{n,m}^{\mathrm{TE}} = -\mathrm{i}$ for $m = 1$, $G_{n,m}^{\mathrm{TE}} = \mathrm{i}$ for

$m = -1$, and $G^{\mathrm{TE}}_{n,m} = 0$ for all other m. The central result of GLMT is that the partial wave scattering and interior amplitudes have the product form (Tam and Corriveau 1978; Gouesbet et al. 1988)

$$\Psi^{\mathrm{scatt,TE}}(r,\theta,\varphi) = -m_2 \sum_{n=1}^{\infty} \sum_{m=-n}^{\infty} C_n \mathbb{B}_{n,m} \zeta_n^{(1)}(m_2 kr) P_n^m[\cos(\theta)] \exp(\mathrm{i}m\varphi) \tag{4.2a}$$

$$\Psi^{\mathrm{inside,TE}}(r,\theta,\varphi) = m_1 \sum_{n=1}^{\infty} \sum_{m=-n}^{n} C_n \mathbb{D}_{n,m} \psi_n(m_1 kr) P_n^m[\cos(\theta)] \exp(\mathrm{i}m\varphi) \tag{4.2b}$$

where

$$\mathbb{B}_{n,m} = b_n G^{\mathrm{TE}}_{n,m}, \tag{4.3a}$$

$$\mathbb{D}_{n,m} = d_n G^{\mathrm{TE}}_{n,m}. \tag{4.3b}$$

The expressions for the TM scalar radiation potentials are similar, but use the beam shape coefficients $G^{\mathrm{TM}}_{n,m}$ multiplied by a_n and c_n.

We now apply this approach to the scattering problem where a single incoming spherical multipole wave in region 2 has partial wavenumber n, azimuthal mode m, and beam shape coefficient $(1/2)G^{\mathrm{TM}}_{n,m}$. The scalar radiation potential of the incident wave and the reflected and transmitted waves is

$$\Psi^{\mathrm{inc,TE}}_{n,m}(r,\theta,\varphi) = (1/2) C_n G^{\mathrm{TE}}_{n,m} \zeta_n^{(2)}(m_2 kr) P_n^m[\cos(\theta)] \exp(\mathrm{i}m\varphi) \tag{4.4a}$$

$$\Psi^{\mathrm{ref,TE}}_{n,m}(r,\theta,\varphi) = C_n \mathbb{R}^{212}_{n.m} \zeta_n^{(1)}(m_2 kr) P_n^m[\cos(\theta)] \exp(\mathrm{i}m\varphi) \tag{4.4b}$$

$$\Psi^{\mathrm{trans,TE}}_{n,m}(r,\theta,\varphi) = C_n \mathbb{T}^{21}_{n,m} \zeta_n^{(2)}(m_1 kr) P_n^m[\cos(\theta)] \exp(\mathrm{i}m\varphi) \tag{4.4c}$$

where

$$\mathbb{R}^{212}_{n.m} = (1/2) R_n^{212} G^{\mathrm{TE}}_{n,m} \tag{4.5a}$$

$$\mathbb{T}^{21}_{n,m} = (1/2) T_n^{21} G^{\mathrm{TE}}_{n,m}. \tag{4.5b}$$

We also apply this approach to another scattering problem where a single outgoing partial wave in region 1 has the partial wave number n, azimuthal mode number m, and beam shape coefficient (1/2) $I^{\mathrm{TE}}_{n,m}$. The scalar radiation potential of the incident wave and the reflected and transmitted waves is

$$\Psi_{n,m}^{\mathrm{inc,TE}}(r,\theta,\varphi) = C_n(1/2)I_{n,m}^{\mathrm{TE}}\zeta_n^{(1)}(m_1kr)P_n^m[\cos(\theta)]\exp(\mathrm{i}m\varphi) \tag{4.6a}$$

$$\Psi_{n,m}^{\mathrm{ref,TE}}(r,\theta,\varphi) = C_n\mathbb{R}_{n,m}^{121}\zeta_n^{(2)}(m_1kr)P_n^m[\cos(\theta)]\exp(\mathrm{i}m\varphi) \tag{4.6b}$$

$$\Psi_{n,m}^{\mathrm{trans,TE}}(r,\theta,\varphi) = C_n\mathbb{T}_{n,m}^{12}\zeta_n^{(1)}(m_2kr)P_n^m[\cos(\theta)]\exp(\mathrm{i}m\varphi), \tag{4.6c}$$

where

$$\mathbb{R}_{n,m}^{121} = (1/2)R_n^{121}I_{n,m}^{\mathrm{TE}} \tag{4.7a}$$

$$\mathbb{T}_{n,m}^{12} = (1/2)T_n^{12}I_{n,m}^{\mathrm{TE}}. \tag{4.7b}$$

Adding Eqs. (4.4c, 4.6a, 4.6b) to obtain the total scalar radiation potential in region 1 we have

$$\begin{aligned}\Psi_{n,m}^{\mathrm{region\,1,TE}}(r,\theta,\varphi) = {} & C_n\left[\mathbb{T}_{n,m}^{21} + (1/2)I_n^{\mathrm{TE}} + \mathbb{R}_{n,m}^{121}\right]\psi_n(m_1kr)P_n^m[\cos(\theta)]\exp(\mathrm{i}m\varphi) \\ & + \mathrm{i}C_n\left[-\mathbb{T}_{n,m}^{21} + (1/2)I_n^{\mathrm{TE}} - \mathbb{R}_{n,m}^{121}\right]\chi_n(m_1kr)P_n^m[\cos(\theta)]\exp(\mathrm{i}m\varphi).\end{aligned} \tag{4.8}$$

In order to have Eq. (4.8) be proportional to $\psi_n(m_1kr)$ alone with no component proportional to $\chi_n(m_1kr)$, as is the case for the interior field of the original Lorenz–Mie scattering problem of Eq. (4.2b), one requires

$$(1/2)I_n^{\mathrm{TE}} = \mathbb{T}_{n,m}^{21} + \mathbb{R}_{n,m}^{121}, \tag{4.9}$$

which, after inserting Eqs. (4.5b, 4.7a) gives

$$I_n^{\mathrm{TE}} = T_n^{21}\left(1 - R_n^{121}\right)^{-1}G_{n,m}^{\mathrm{TE}}. \tag{4.10}$$

Substituting Eq. (4.9) into Eq. (4.8) for the total interior field and equating the result to Eq. (4.2b) for the Lorenz–Mie interior field, we find that

$$(m_1/m_2)\mathbb{D}_{n,m} = 2(\mathbb{T}_{n,m}^{21} + \mathbb{R}_{n,m}^{121}). \tag{4.11}$$

Next, substituting Eqs. (4.3b, 4.5b, 4.7a) into Eq. (4.11), and then using Eq. (4.10) to eliminate the dependence on $I_{n,m}^{\mathrm{TE}}$, one obtains

$$\begin{aligned}(m_1/m_2)d_nG_{n,m}^{\mathrm{TE}} &= \left[T_n^{21} + R_n^{121}\left(1-R_n^{121}\right)^{-1}T_n^{21}\right]G_{n,m}^{\mathrm{TE}} \\ &= \left(1-R_n^{121}\right)^{-1}T_n^{21}G_{n,m}^{\mathrm{TE}}.\end{aligned} \tag{4.12}$$

Since this is true for every set of beam shape coefficients, the Debye series decomposition of the partial wave interior amplitude of Lorenz–Mie theory is (Lock 1988)

$$d_n = (m_2/m_1)T_n^{21}\left(1-R_n^{121}\right)^{-1}. \tag{4.13}$$

In order to obtain the Debye series for b_n, we add Eqs. (4.4a, 4.4b, 4.6c) to obtain the total scalar radiation potential in region 2. Then doing a little bit of rearranging we obtain

$$\begin{aligned}\Psi_{n,m}^{\text{region 2,TE}}(r,\theta,\varphi) = {} & C_n G_{n,m}^{\text{TE}}\psi_n(m_2kr)P_n^m[\cos(\theta)]\exp(\mathrm{i}m\varphi) \\ & -C_n\left[(1/2)G_{n,m}^{\text{TE}} - \mathbb{R}_{n,m}^{212} - \mathbb{T}_{n,m}^{12}\right]\zeta_n^{(1)}(m_2kr)P_n^m[\cos(\theta)]\exp(\mathrm{i}m\varphi).\end{aligned} \tag{4.14}$$

This is identical in form to the total exterior scalar radiation potential of Eqs. (4.1, 4.2a) for the original generalized Lorenz–Mie scattering problem. Equating (4.14) to Eqs. (4.1, 4.2a), we find that

$$\mathbb{B}_{n,m} = (1/2)G_{n,m}^{\text{TE}} - \mathbb{R}_{n,m}^{212} - \mathbb{T}_{n,m}^{12}. \tag{4.15}$$

Lastly, substituting Eqs. (4.3a, 4.5a, 4.7b) into Eq. (4.15), and then substituting Eq. (4.10) into the result in order to eliminate the dependence on $I_{n,m}^{\text{TE}}$, one obtains

$$b_n G_{n,m}^{\text{TE}} = (1/2)\left[1-R_n^{212}-T_n^{21}\left(1-R_n^{121}\right)^{-1}T_n^{12}\right]G_{n,m}^{\text{TE}}. \tag{4.16}$$

Since this is again true for every set of beam shape coefficients, the Debye series decomposition of the Lorenz–Mie partial wave scattering amplitude is (van der Pol and Bremmer 1937; Franz 1954; Rubinow 1961; Chen 1964; Nussenzveig 1969a)

$$b_n = (1/2)\left[1-R_n^{212}-T_n^{21}\left(1-R_n^{121}\right)^{-1}T_n^{12}\right]. \tag{4.17}$$

The derivation for the TM polarization proceeds exactly the same way.

In order to make the physical interpretation of each term of the Debye series more evident, we note that the partial wave reflection amplitude R_n^{121} satisfies $\left|R_n^{121}\right|<1$, so that the reciprocal term in Eqs. (4.13, 4.17) may be expanded as a geometric series, giving

$$\begin{aligned}a_n, b_n &= (1/2)\left[1-R_n^{212} - \sum_{p=1}^{\infty} T_n^{21}\left(R_n^{121}\right)^{p-1}T_n^{12}\right] \\ &= (1/2)\left[1-R_n^{212} - (m_1/m_2)(c_n,d_n)T_n^{12}\right]\end{aligned} \tag{4.18a}$$

$$c_n, d_n = \sum_{p=1}^{\infty} (m_2/m_1) T_n^{21} \left(R_n^{121}\right)^{p-1}. \tag{4.18b}$$

These expressions are not an approximation to the partial wave scattering and interior amplitudes; they are exact. As to notation, just as the superscripts of the R and T factors read from left to right, the ordering of the factors also reads from left to right as one follows the progression of the partial wave through the sphere. The first term of Eq. (4.18a), when substituted into Eqs. (2.11a, 2.11b) for a_n and b_n and then summed over partial waves, will be seen in Sect. 5 to be interpreted as the far-zone fields resulting from diffraction of the incident plane wave by the sphere. The second term is interpreted as the externally reflected partial waves, and the terms in the sum over p are interpreted as transmission of the partial waves into the sphere followed by $p - 1$ internal reflections, and then transmission back out. These terms are pictorially illustrated in Fig. 2. There is a technical caveat that needs to be mentioned concerning this interpretation. Neither the partial wave sum of diffraction nor of external reflection alone is convergent in the $n \to \infty$ limit. However, the sum of the two terms taken together is convergent (Nussenzveig 1969a). Thus diffraction-plus-external reflection is standardly taken to be the $p = 0$ term of the Debye series.

In order to illustrate how the Debye series provides a physical interpretation of the complicated structure observed in the Lorenz–Mie scattered intensity in the $a \gg \lambda$ limit, Fig. 3 plots the individual Debye contributions to the scattered intensity for $0 \le p \le 3$ and $\lambda = 0.65$ μm, $a = 10$ μm, $m_1 = 1.33257$, and $m_2 = 1$, which were the parameters used in Fig. 1. The diffraction portion of the $p = 0$ term is seen to dominate for $0° \le \theta \le 2°$. Thereafter, interference between $p = 1$ transmission and the external reflection portion of $p = 0$ dominates for $2° \le \theta \le 80°$, except in the vicinity of the TM external reflection Brewster angle at $\theta^B = 73.77°$ [see Sect. 4.3.2 of Hecht (1987)]. The $p = 3$ second order rainbow at $\theta^R = 129.22°$ along with its supernumeraries interfere with external reflection and dominate TE scattering for $110° \le \theta \le 130°$, while the brighter $p = 2$ first-order rainbow at $\theta^R = 137.86°$ along with its supernumeraries dominate TE scattering for $135° \le \theta \le 170°$. Strong rainbow enhancements in the scattered intensity do not occur for TM scattering since the internal reflections involved at $\theta_1^R = 40.25°$ for $p = 2$ and $\theta_1^R = 45.49°$ for $p = 3$ occur near the Brewster angle of $\theta_1^B = 36.89°$. Figure 4 compares the Lorenz–Mie scattered intensity of Fig. 1 with the intensity corresponding to the $0 \le p \le 7$ terms of the Debye series. The results are virtually identical, as is expected since the entire Debye series is an exact representation of the Lorenz–Mie partial wave scattering amplitudes.

Lastly, for comparison with the results of Sects. 8 and 9 it should be noted that for a scalar plane wave incident on a homogeneous sphere, as occurs for scattering of sound waves or quantum mechanical waves, the incident, scattered, and interior wave functions are

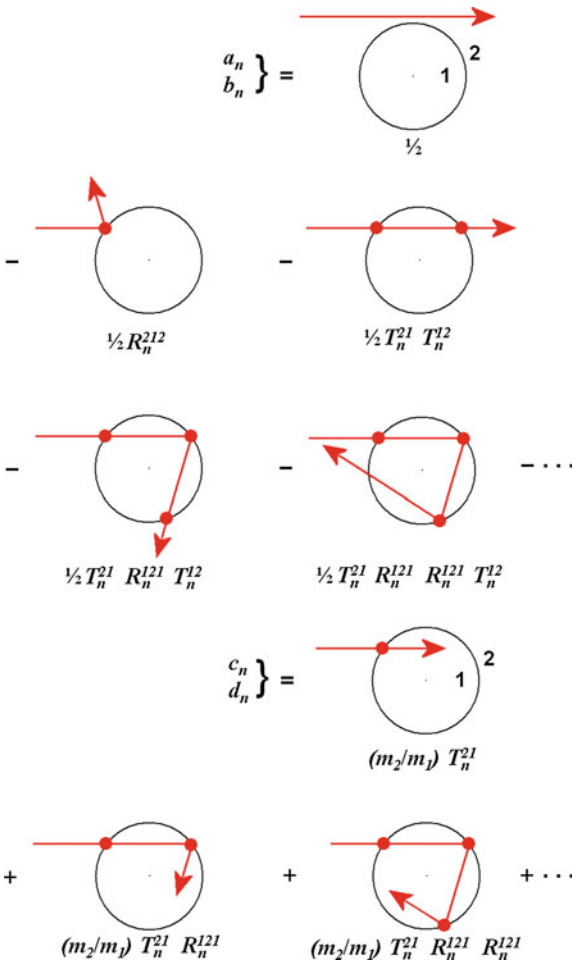


Fig. 2 Terms of the Debye series expansion of the partial wave scattering amplitudes a_n, b_n, and the partial wave interior amplitudes c_n, d_n

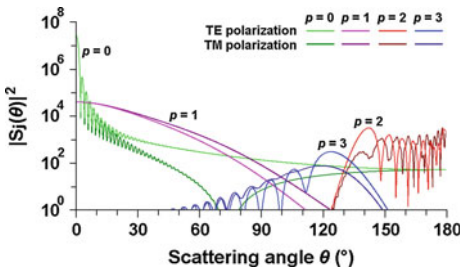


Fig. 3 Debye series terms for $0 \le p \le 3$ of the TE and TM intensity as a function of angle for scattering of an x-polarized plane wave by a sphere with $\lambda = 0.65$ μm, $a = 10$ μm, $m_1 = 1.33257$, and $m_2 = 1$

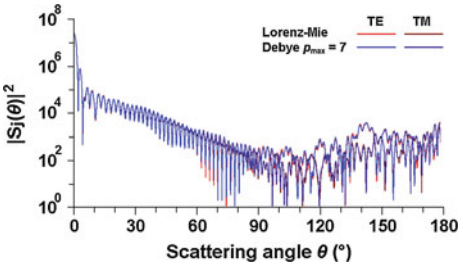


Fig. 4 Sum of the Debye series terms for $0 \le p \le 7$ of the TE and TM intensity as a function of angle for scattering of an x-polarized plane wave by a sphere with $\lambda = 0.65$ μm, $a = 10$ μm, $m_1 = 1.33257$, and $m_2 = 1$

$$\psi^{\text{inc}}(r,\theta,\varphi) = m_2 E_0 \sum_{n=0}^{\infty} j_n(m_2 kr) P_n^m[\cos(\theta)]\exp(\mathrm{i}m\varphi) \tag{4.19a}$$

$$\psi^{\text{scatt}}(r,\theta,\varphi) = -m_2 E_0 \sum_{n=0}^{\infty} b_n h_n^{(1)}(m_2 kr) P_n^m[\cos(\theta)]\exp(\mathrm{i}m\varphi) \tag{4.19b}$$

$$\psi^{\text{inside}}(r,\theta,\varphi) = m_1 E_0 \sum_{n=0}^{\infty} d_n j_n(m_1 kr) P_n^m[\cos(\theta)]\exp(\mathrm{i}m\varphi), \tag{4.19c}$$

and the boundary conditions at the surface of the sphere are the continuity of ψ and $\partial\psi/\partial r$. The resulting partial wave scattering and interior amplitudes are identical to the TE version of Eqs. (2.18, 2.21), and the four fundamental scattering amplitudes N_n, D_n, P_n, and Q_n are identical to the TE version of Eqs. (2.15a–2.15d) divided by $m_1 m_2 (ka)^2$. The Debye series expansion of b_n and d_n is identical to Eqs. (4.18a, 4.18b).

5 Derivation of the Ray Model of Light Scattering Using the Debye Series

The geometrical ray model of light scattering provides a simple and intuitive description of the interaction of an incident beam with a target particle in the short wavelength limit. As was mentioned in Sect. 1, the Debye series form of the partial wave scattering amplitudes can be used to provide a rigorous derivation of ray theory since Lorenz–Mie theory is the exact solution to the electromagnetic boundary value problem of a plane wave scattered by a homogeneous spherical particle, and the Debye series is an exact representation of the partial wave scattering amplitudes of Lorenz–Mie theory (van de Hulst 1981; Debye 1908). The derivation is outlined as follows. The first thing one must do is motivate van de Hulst's localization principle which associates a geometrical ray with a partial wave of the incident beam [see pp. 208–209 of van de Hulst (1981)]. Consider the semiclassical regime $a \gg \lambda$ where many partial waves are required for convergence

of the partial wave sums. In quantum mechanics the orbital angular momentum of a physical system, here to taken to be the partial wave n incident on the scattering particle whose center is at the origin of coordinates, is

$$L = [n(n+1)]^{1/2}\hbar \approx (n+1/2)\hbar \tag{5.1}$$

where $\hbar$ is Planck's constant [see Sect. 4.3 of Griffiths (2005)]. In classical mechanics, the orbital angular momentum of a projectile incident on a target particle of radius a whose center is at the origin is

$$L = pa\sin(\theta_2) \tag{5.2}$$

where p is the projectile momentum (not to be confused with the integer index p of the Debye series) and θ_2 is the angle of incidence of the projectile on the target as shown in Fig. 5.

According to the de Broglie relations for a photon in a material of refractive index m_2, one has the association [see p. 19 of Griffiths (2005)]

$$p = m_2\hbar k. \tag{5.3}$$

Mixing together these different classical and quantum mechanical expressions for orbital angular momentum suggests that in the semiclassical regime the partial wave n can be associated with the impact parameter

$$b = a\sin(\theta_2) \tag{5.4}$$

of a light ray via

$$n + 1/2 \approx m_2 kb \tag{5.5}$$

with $b \leq a$. Since the incident plane wave has infinite lateral extent, the sum over partial waves $1 \leq n < \infty$ can be qualitatively associated with an infinite collection of

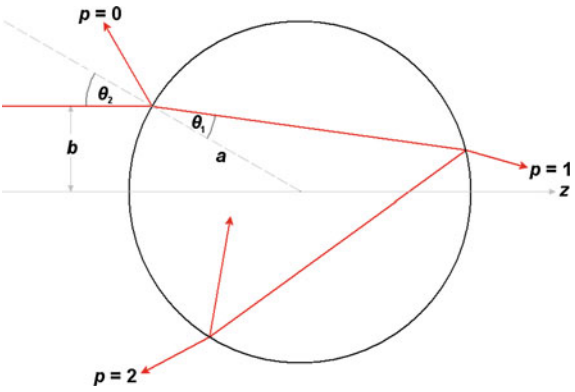


Fig. 5 Geometry of an incident ray with impact parameter b interacting with a spherical particle of radius a

parallel incident light rays, all having different impact parameters with respect to the z axis. Light rays with $b \leq a$ (or partial waves with $n \leq x$) directly interact with the sphere, while those with $b > a$ (or partial waves with $n > x$) pass the sphere by without striking it and do not directly interact with it. This interpretation validates the first term of Eq. (2.13) for the highest partial wave contribution $n_{\max}$ to the scattering amplitudes in the sum over partial waves. The term in Eq. (2.13) proportional to $x^{1/3}$ is a wave theory correction describing partial waves that tunnel through the centrifugal barrier surrounding the sphere before interacting with it [see Sect. 8.1 of Nussenzveig (1992)].

We next use this association to show that the first term of the Debye series of Eq. (4.18a) describes diffraction of the incident plane wave [see pp. 209–210 of van de Hulst (1981)]. Substituting the first term of Eq. (4.18a) for a_n and b_n into Eqs. (2.11a, 2.11b), one obtains

$$S_1^{\text{diff}}(\theta) = S_2^{\text{diff}}(\theta) = \sum_{n=1}^{n_{\max}} \{(2n+1)/[2n(n+1)]\}[\pi_n(\theta) + \tau_n(\theta)]$$
$$\approx \sum_{n=1}^{n_{\max}} [1/(n+1/2)][\pi_n(\theta) + \tau_n(\theta)]. \tag{5.6}$$

Since diffraction is expected to be most important in the near-forward direction where azimuthal focusing of the scattered light is important, we must evaluate the Lorenz–Mie angular functions $\pi_n(\theta)$ and $\tau_n(\theta)$ for small angles where $\sin(\theta) \approx \theta$ and $\cos(\theta) \approx 1$. In this limit, the associated Legendre differential equation becomes approximately equal to Bessel's differential equation [see Eq. (9.1.1) of Abramowitz and Stegun (1964)]. This results in $P_n^m[\cos(\theta)]$ being proportional to the Bessel function $J_m[(n + ½)\theta]$. Using the known value of $\pi_n(0)$ and $\tau_n(0)$ to evaluate the constant of proportionality, one obtains [see p. 209 of van de Hulst (1981)]

$$\pi_n(\theta) \approx [(n+1/2)^2/2]\{J_0[(n+1/2)\theta] + J_2[(n+1/2)\theta]\} \tag{5.7a}$$

$$\tau_n(\theta) \approx [(n+1/2)^2/2]\{J_0[(n+1/2)\theta] - J_2[(n+1/2)\theta]\}. \tag{5.7b}$$

Then approximating the sum over partial waves by an integral over the associated impact parameter of Eq. (5.5) (Ford and Wheeler 1959; Berry and Mount 1972), Eq. (5.6) simplifies to

$$S_1^{\text{diff}}(\theta) = S_2^{\text{diff}}(\theta) = \sum_{n=1}^{m_2 ka} (n+1/2) J_0[(n+1/2)\theta]$$
$$\approx \int_0^{m_2 ka} (m_2 k)^2 b \, \mathrm{d}b J_0(m_2 k b \theta) = (m_2 ka)^2 [J_1(m_2 ka\theta)/(m_2 ka\theta)]. \tag{5.8}$$

The far-zone diffracted portion of the electric and magnetic fields of Eqs. (2.10a, 2.10b) are then

$$\mathbf{E}^{\mathbf{diff}} \approx \left[\mathrm{i}(m_2 ka)^2 E_0 \exp(\mathrm{i}m_2 kr)/(m_2 kr)\right][J_1(m_2 ka\theta)/(m_2 ka\theta)]\mathbf{u_x} \tag{5.9a}$$

$$(c/m_2)\mathbf{B}^{\mathbf{diff}} \approx \left[\mathrm{i}(m_2 ka)^2 E_0 \exp(\mathrm{i}m_2 kr)/(m_2 kr)\right][J_1(m_2 ka\theta)/(m_2 ka\theta)]\mathbf{u_y}, \tag{5.9b}$$

and are exactly what one would expect for Fraunhofer diffraction of a plane wave in medium 2 past a circular obstacle of radius a.

Since the derivation of the ray theory scattered fields is similar for all values of $p \geq 1$, we describe only the derivation for the $p = 2$ term for scattering by a particle with $m_1 > m_2$, where

$$a_n, b_n = -(1/2)T_n^{21} R_n^{121} T_n^{12}. \tag{5.10}$$

The derivation proceeds in four steps. The first step consists of approximating the partial wave transmission and reflection amplitudes in the $a \gg \lambda$ limit. The Riccati–Hankel functions and their derivatives appearing in Eqs. (3.5a, 3.5b, 3.8) can be approximated using Debye's asymptotic form for $n < m_j ka$ [i.e., on the oscillatory side of the transition region centered on $n = m_j ka$, see Sect. 9.3 of Abramowitz and Stegun (1964)]. One has

$$\zeta_n^{(1)}(z) \approx C(n,z)^{-1/2} \exp[\mathrm{i}\Phi_n(z)] \tag{5.11a}$$

$$\zeta_n^{(1)\prime}(z) \approx \mathrm{i}C(n,z)^{1/2} \exp[\mathrm{i}\Phi_n(z)] \tag{5.11b}$$

$$\zeta_n^{(2)}(z) \approx C(n,z)^{-1/2} \exp[-\mathrm{i}\Phi_n(z)] \tag{5.11c}$$

$$\zeta_n^{(2)\prime}(z) \approx -\mathrm{i}C(n,z)^{1/2} \exp[-\mathrm{i}\Phi_n(z)], \tag{5.11d}$$

where

$$C(n,z) = \left\{1 - [(n+1/2)/z]^2\right\}^{1/2} \tag{5.12}$$

and

$$\Phi_n(z) \approx z\,C(n,z) - (n+1/2)\arcsin[C(n,z)] - \pi/4. \tag{5.13}$$

This result can be derived by applying the WKB approximation to the spherical Bessel function differential equation in the vicinity of the classical turning point and making use of $n(n+1) \approx (n+1/2)^2$ for $n \gg 1$. Using the heuristic association of

partial waves with impact parameters of Eq. (5.5) and the geometry of Fig. 5, one obtains

$$C(n, m_2 ka) = \cos(\theta_2), \tag{5.14}$$

and using Snell's law of refraction

$$m_2 \sin(\theta_2) = m_1 \sin(\theta_1) \tag{5.15}$$

at the surface of the sphere, one has

$$C(n, m_1 ka) = \cos(\theta_1). \tag{5.16}$$

Substituting Eqs. (5.10–5.16) into Eqs. (3.5a, 3.5b, 3.8), the partial wave transmission and reflection amplitudes become approximately

$$T_n^{21} \approx 2m_1 \exp[i\Phi_n(m_1 ka) - i\Phi_n(m_2 ka)][\cos(\theta_1)\cos(\theta_2)]^{1/2}/[\alpha \cos(\theta_1) + \beta \cos(\theta_2)] \tag{5.17a}$$

$$T_n^{12} \approx 2m_2 \exp[i\Phi_n(m_1 ka) - i\Phi_n(m_2 ka)][\cos(\theta_1)\cos(\theta_2)]^{1/2}/[\alpha \cos(\theta_1) + \beta \cos(\theta_2)] \tag{5.17b}$$

$$R_n^{212} \approx -\exp[-2i\Phi_n(m_2 ka)][\alpha \cos(\theta_1) - \beta \cos(\theta_2)]/[\alpha \cos(\theta_1) + \beta \cos(\theta_2)] \tag{5.17c}$$

$$R_n^{121} \approx \exp[2i\Phi_n(m_1 ka)][\alpha \cos(\theta_1) - \beta \cos(\theta_2)]/[\alpha \cos(\theta_1) + \beta \cos(\theta_2)]. \tag{5.17d}$$

The magnitude of these expressions bears a great similarity to the TE and TM Fresnel coefficients for the transmitted and reflected amplitude of a plane wave diagonally incident on a flat interface separating regions 1 and 2 [see Sect. 4.3.2 of Hecht (1987)],

$$t^{21}(\theta_2) = 2m_2 \cos(\theta_2)/[\alpha \cos(\theta_1) + \beta \cos(\theta_2)] \tag{5.18a}$$

$$t^{12}(\theta_2) = 2m_1 \cos(\theta_1)/[\alpha \cos(\theta_1) + \beta \cos(\theta_2)] \tag{5.18b}$$

$$r^{212}(\theta_2) = -[\alpha \cos(\theta_1) - \beta \cos(\theta_2)]/[\alpha \cos(\theta_1) + \beta \cos(\theta_2)] \tag{5.18c}$$

$$r^{121}(\theta_2) = [\alpha \cos(\theta_1) - \beta \cos(\theta_2)]/[\alpha \cos(\theta_1) + \beta \cos(\theta_2)]. \tag{5.18d}$$

The magnitude of R_n^{212} and R_n^{121} exactly matches that of r^{212} and r^{121}. Although the magnitude of T_n^{21} and T_n^{12} does not exactly match that of t^{21} and t^{12}, the magnitude of the product $T_n^{21} T_n^{12}$, which appears in Eq. (5.10), does exactly match. The reason for this similarity is that an incoming light ray with $1 \ll n \ll m_2 ka$ interacts with a

sufficiently small portion of the surface of the scattering particle that it cannot detect its local curvature. Rather, the surface of the particle looks tilted and locally flat to the incident ray.

The second step of the derivation is the simplification of the partial wave angular functions $\pi_n(\theta)$ and $\tau_n(\theta)$ for $n \gg 1$ in the angular interval sufficiently far from forward scattering and back-scattering focusing, $0^\circ \ll \theta \ll 180^\circ$. This approximation is obtained by: (i) solving Legendre's differential equation for $P_n[\cos(\theta)]$ at $\theta = 90^\circ$ and using $n(n + 1) \approx (n + 1/2)^2$, (ii) using the known value of $P_n(0)$ [see Eqs. (12.34–12.35) of Arfken (1985)] to determine the amplitude of the resulting sinusoidal function, (iii) using Sterling's approximation to eliminate the factorials obtained in step (ii), and then (iv) substituting the result back into Legendre's equation to determine the residual θ dependence. This procedure gives [see Eq. (8.10.7) of Abramowitz and Stegun (1964)]

$$P_n[\cos(\theta)] \approx \{2/[(n+1/2)\pi \sin(\theta)]\}^{1/2}\cos[(n+1/2)\theta-\pi/4]. \tag{5.19}$$

Then using Eqs. (2.12a, 2.12b), one obtains [see p. 212 of van de Hulst (1981)]

$$\pi_n(\theta) \approx \left\{[2(n+1/2)]/\left[\pi \sin^3(\theta)\right]\right\}^{1/2}\sin[(n+1/2)\theta-\pi/4] \tag{5.20a}$$

$$\tau_n(\theta) \approx \left\{\left[2(n+1/2)^3\right]/[\pi \sin(\theta)]\right\}^{1/2}\cos[(n+1/2)\theta-\pi/4]. \tag{5.20b}$$

Since $\tau_n(\theta)$ with its factor of $(n + 1/2)^{3/2} \sin^{-1/2}(\theta)$ is generally much larger for $n \gg 1$ than $\pi_n(\theta)$ with its factor of $(n + 1/2)^{1/2} \sin^{-3/2}(\theta)$, the scattering amplitudes of Eqs. (2.11a, 2.11b) become approximately

$$S_1(\theta) \approx \sum_{n=1}^{\infty}\{(2n+1)/[n(n+1)]\}b_n\tau_n(\theta) \tag{5.21a}$$

$$S_2(\theta) \approx \sum_{n=1}^{\infty}\{(2n+1)/[n(n+1)]\}a_n\tau_n(\theta). \tag{5.21b}$$

As was mentioned in Sect. 1, the function $S_1(\theta)$ is associated with the TE scattering amplitude of geometrical light rays, and $S_2(\theta)$ is associated with the TM scattering amplitude. Substituting Eqs. (5.10, 5.17a, 5.17b, 5.17d, 5.20a, 5.20b) into Eq. (5.21a) for the TE amplitude, one obtains

$$S_1^{p=2}(\theta) \approx -\{1/[2\pi \sin(\theta)]\}^{1/2}\sum_{n=1}^{\infty}(n+1/2)^{1/2}A_n\left\{\exp\left[i\left(\Omega_n^+ -\pi/4\right)\right] + \exp\left[i\left(\Omega_n^- +\pi/4\right)\right]\right\} \tag{5.22}$$

where

$$A_n = 4m_1 m_2 \cos(\theta_1)\cos(\theta_2)/[m_1\cos(\theta_1) + m_2\cos(\theta_2)]^3 \tag{5.23}$$

with θ_1 and θ_2 being functions of n as in Eqs. (5.14, 5.16), and

$$\Omega_n^{\pm} = 4\Phi_n(m_1 ka) - 2\Phi_n(m_2 ka) \pm (n + 1/2)\theta. \tag{5.24}$$

The third step in the derivation is approximately converting the sum over partial waves into an integral over an associated impact parameter. Letting the continuous variable u be associated with the partial wave number n via

$$u = n + 1/2 = m_2 ka\sin(\theta_2), \tag{5.25}$$

the sum over partial waves can be approximated by (Ford and Wheeler 1959; Berry and Mount 1972)

$$S_1^{p=2}(\theta) \approx -\{1/[2\pi\sin(\theta)]\}^{1/2}\int_0^{\infty} du\, u^{1/2}A(u)\{\exp[i\Gamma^+(u) - 3i\pi/4] + \exp[i\Gamma^-(u) - i\pi/4]\} \tag{5.26}$$

with

$$\begin{aligned}\Gamma^{\pm}(u) = {} & 4m_1 ka\left[1-(u/m_1 ka)^2\right]^{1/2} - 4u\arcsin\left[1-(u/m_1 ka)^2\right]^{1/2} \\ & - 2m_2 ka\left[1-(u/m_2 ka)^2\right]^{1/2} + 2u\arcsin\left[1-(u/m_2 ka)^2\right]^{1/2} \pm u\theta.\end{aligned} \tag{5.27}$$

The final step in the derivation is evaluating the integral using the stationary phase approximation [see Sect. 4.2a of Felsen and Marcuvitz (1973)]. The functions $u^{1/2}$ and $A(u)$ of Eq. (5.26) are slowly varying in u, while the phase functions $\Gamma^{\pm}(u)$ are rapidly varying. If either phase function possesses a relative maximum or minimum, it corresponds to a region of constructive interference of the phase when the impact parameter u is integrated, resulting in a large contribution to the integral. If there is no relative maximum or minimum of the phase, only destructive interference will occur when u is integrated over, contributing negligibly to the value of the integral.

Let a positive angle θ correspond to a light ray incident on the sphere above the z axis and deflected below the z axis as in Fig. 5. Then for $p = 2$ one has

$$d\Gamma^{\pm}/du = -(\pi + 2\theta_2 - 4\theta_1) \pm \theta. \tag{5.28}$$

The phase function $\Gamma^{+}(u)$ has a stationary point at

$$\theta = \pi + 2\theta_2 - 4\theta_1, \tag{5.29}$$

while $\Gamma^{-}(u)$ has no stationary point. Equation (5.26) for TE scattering then becomes approximately

$$S_1^{p=2}(\theta) \approx -\exp(-3\mathrm{i}\pi/4)\{1/[2\pi\,\sin(\theta)]\}^{1/2} u_{\mathrm{sp}}^{1/2} A(u_{\mathrm{sp}}) \int_0^{\infty} \mathrm{d}u\, \exp[i\Gamma^{+}(u)], \tag{5.30}$$

where $u_{\mathrm{sp}} = m_2 ka\,\sin(\theta_2)$ is the stationary point of u consistent with Eq. (5.29), and the slowly varying functions $A(u)$ and $u^{1/2}$ are to be evaluated at that stationary point. As θ_2 increases from 0° for head-on incidence to 90° for grazing incidence when u is integrated over, the minimum value of θ of Eq. (5.29) occurs when $\mathrm{d}\theta/\mathrm{d}\theta_2 = 0$, or

$$\cos(\theta_2^R) = \left\{\left[(m_1/m_2)^2 - 1\right]/3\right\}^{1/2}. \tag{5.31}$$

The angle θ_2^R substituted into Eqs. (5.15, 5.29) gives the Descartes scattering angle θ^{R} of the first-order rainbow in ray theory (Walker 1976). The derivative $\mathrm{d}\theta/\mathrm{d}\theta_2$ turns out to be negative for $0° \le \theta < \theta_2^{\mathrm{R}}$, and it is positive for $\theta_2^R < \theta \le 90°$. In addition

$$\mathrm{d}^2\Gamma^{+}/\mathrm{d}u^2 = -\{1/[m_2 ka\cos(\theta_2)]\}\mathrm{d}\theta/\mathrm{d}\theta_2. \tag{5.32}$$

The sign of $\mathrm{d}^2\Gamma^{+}/\mathrm{d}u^2$ determines an additional phase factor in Eq. (5.34) below when the stationary phase method is applied. The stationary phase evaluation of the integral in Eq. (5.30) for $\theta_2 \neq \theta_2^R$ gives

$$\begin{aligned} S_1^{p=2}(\theta) \approx\ & m_2 ka\{\sin(\theta_2)\cos(\theta_2)/[2\,\sin(\theta)]\}^{1/2} |\{[2m_2\,\cos(\theta_2)]/[m_1\,\cos(\theta_1)]\} - 1|^{-1/2} \\ & \times t^{21}(\theta_2) r^{121}(\theta_2) t^{12}(\theta_2) \exp\{\mathrm{i}[4m_1 ka\,\cos(\theta_1) - 2m_2 ka\,\cos(\theta_2)]\} \exp(\mathrm{i}\Phi_{\mathrm{npl}}), \end{aligned} \tag{5.33}$$

where the non-path-length phase is given by

$$\begin{aligned} \Phi_{\mathrm{npl}} &= \pi/2 \quad \text{for } 0° \le \theta_2 < \theta_2^R \\ &= 0 \quad \text{for } \theta_2^R < \theta_2 \le 90°. \end{aligned} \tag{5.34}$$

The first complex exponential in Eq. (5.33) is the accumulated phase of a ray's path through the sphere. The second complex exponential is van de Hulst's non-path length phase, which decreases by 90° as the Descartes angle of the rainbow caustic

is crossed [see p. 207 of van de Hulst (1981)]. The $p = 2$ ray theory scattered electric and magnetic fields are obtained by substituting Eq. (5.33) into Eqs. (2.10a-b). The factor to the $-1/2$ power in Eq. (5.33) diverges at the Descartes scattering angle $\theta = \theta_2^R$, giving an infinite intensity at the first-order rainbow in ray theory. However, this value of θ lies outside the region of validity of the stationary phase approximation that was used to obtain Eq. (5.33). Since $d^2\Gamma^+/du^2 = 0$ there, one must consider the third order term in the Taylor series expansion of $\Gamma^+(u)$. This leads to the integral of Eq. (5.30) becoming an Airy integral [see Sect. 10.4 of Abramowitz and Stegun (1964)], which results in the scattered fields being large but finite in the vicinity of the Descartes angle (i.e., the Airy shift). The fields decreases faster than exponentially on one side of the rainbow peak (i.e., the complex ray regime), and are oscillatory on the other side of the peak (i.e., the supernumerary ray interference pattern).

6 Scattering of a Neumann Beam by a Homogeneous Sphere

It was seen in Eqs. (2.18, 2.21) that the compact form of the partial wave scattering and interior amplitudes a_n, b_n, c_n, and d_n for plane wave incidence depended only on two of the four fundamental partial wave amplitudes of Eqs. (2.15a–2.15d), N_n and D_n. This is also the case for a large class of beams in GLMT such as focused Gaussian beams, Bessel beams, top-hat beams, etc. Since there are four fundamental partial wave amplitudes, N_n, D_n, P_n, and Q_n, it could be asked if there is a type of incident beam for which the compact form of the partial wave scattering and interior amplitudes depends only on the other two fundamental partial wave amplitudes P_n and Q_n? This section considers this question.

Consider a TE-polarized beam whose scalar radiation potential in region 2 is

$$\Psi^{\mathrm{inc,TE}}(r,\theta,\varphi) = -\mathrm{i}m_2E_0\sum_{n=1}^{\infty} C_n\chi_n(m_2kr)P_n^1[\cos(\theta)]\sin(\varphi). \tag{6.1}$$

This has been called a Neumann beam (Lock 2011) since the radial dependence is a Riccati–Neumann function rather than a Riccati–Bessel function. In the absence of a scattering particle this beam has an essential singularity at the origin. Just as $\zeta_n^{(1)}$ and $\zeta_n^{(2)}$ with the time dependence $\exp(-\mathrm{i}\omega t)$ are radially outgoing and incoming traveling waves, ψ_n and χ_n can be thought of as two orthogonal radial standing waves, with ψ_n being regular at the origin and χ_n diverging there. The Neumann beam of Eq. (6.1) can be interpreted as the standing wave formed by a radially outgoing wave from a point source located at the center of a large perfectly reflecting spherical shell and the radially incoming wave that has been totally reflected back toward the origin by the shell.

It was shown in Lock (2011) that this standing wave leads to scattered and interior fields whose compact form depends only on P_n and Q_n. The scattered and interior scalar radiation potentials for this scattering problem are

$$\Psi^{\text{scatt,TE}}(r,\theta,\varphi) = -m_2 E_0 \sum_{n=1}^{\infty} C_n b_n \zeta_n^{(1)}(m_2 kr) P_n^1[\cos(\theta)]\sin(\varphi) \tag{6.2a}$$

$$\Psi^{\text{inside,TE}}(r,\theta,\varphi) = -\mathrm{i}m_1 E_0 \sum_{n=1}^{\infty} C_n d_n \chi_n(m_1 kr) P_n^1[\cos(\theta)]\sin(\varphi). \tag{6.2b}$$

Matching boundary conditions for the continuity of the tangential components of the electric and magnetic fields at the surface of the sphere gives

$$a_n, b_n = -Q_n/(Q_n - \mathrm{i}P_n) \tag{6.3a}$$

$$c_n, d_n = -\mathrm{i}m_2/(Q_n - \mathrm{i}P_n) \tag{6.3b}$$

as desired. Using the procedure of Sects. 3 and 4 to carry out the Debye series derivation gives

$$a_n, b_n = (1/2)\left[-1 - R_n^{212} + \sum_{p=1}^{\infty} T_n^{21}\left(-R_n^{121}\right)^{p-1} T_n^{12}\right] \tag{6.4a}$$

$$c_n, d_n = (m_2/m_1)\sum_{p=1}^{\infty} T_n^{21}\left(-R_n^{121}\right)^{p-1}. \tag{6.4b}$$

in analogy to Eqs. (4.18a, 4.18b) as the expanded form of the partial wave scattering and interior amplitudes.

7 Debye Series for a Scattering by a Coated Sphere and a Multi-layer Sphere

7.1 Scattering by a Coated Sphere

Consider now a spherical particle of refractive index m_1 and radius a_{12} (region 1) concentrically surrounded by a material of refractive index m_2 (region 2) in an exterior medium of refractive index m_3 (region 3). The overall radius of the coated sphere is a_{23}. An incident linearly polarized electromagnetic plane wave is scattered by the coated sphere [see Aden and Kerker (1951) and Sect. 8.1 of Bohren and Huffman (1983)]. The partial wave scattering amplitudes a_n and b_n, as well as the partial wave amplitudes c_n and d_n in the core and $K_n^{\pm}$ and $L_n^{\pm}$ in the coating, where

$\pm$ indicates radially outgoing and incoming waves, can be written in a number of different but equivalent ways. Perhaps the simplest way, which turns out to anticipate the Debye series expansion of these amplitudes, is as follows (Lock 2005). In analogy to Eqs. (2.15a–2.15d) for a homogeneous sphere, the four fundamental partial wave amplitudes for the 12 interface separating regions 1 and 2 are

$$N_n^{12} = \alpha\,\psi_n(m_2ka_{12})\psi_n'(m_1ka_{12}) - \beta\,\psi_n'(m_2ka_{12})\psi_n(m_1ka_{12}) \tag{7.1a}$$

$$D_n^{12} = \alpha\,\chi_n(m_2ka_{12})\psi_n'(m_1ka_{12}) - \beta\,\chi_n'(m_2ka_{12})\psi_n(m_1ka_{12}) \tag{7.1b}$$

$$P_n^{12} = \alpha\,\psi_n(m_2ka_{12})\chi_n'(m_1ka_{12}) - \beta\,\psi_n'(m_2ka_{12})\chi_n(m_1ka_{12}) \tag{7.1c}$$

$$Q_n^{12} = \alpha\,\chi_n(m_2ka_{12})\chi_n'(m_1ka_{12}) - \beta\,\chi_n'(m_2ka_{12})\chi_n(m_1ka_{12}), \tag{7.1d}$$

and their counterparts for the 23 interface separating regions 2 and 3 are

$$N_n^{23} = \gamma\,\psi_n(m_3ka_{23})\psi_n'(m_2ka_{23}) - \delta\,\psi_n'(m_3ka_{23})\psi_n(m_2ka_{23}) \tag{7.2a}$$

$$D_n^{23} = \gamma\,\chi_n(m_3ka_{23})\psi_n'(m_2ka_{23}) - \delta\,\chi_n'(m_3ka_{23})\psi_n(m_2ka_{23}) \tag{7.2b}$$

$$P_n^{23} = \gamma\,\psi_n(m_3ka_{23})\chi_n'(m_2ka_{23}) - \delta\psi_n'(m_3ka_{23})\chi_n(m_2ka_{23}) \tag{7.2c}$$

$$Q_n^{23} = \gamma\,\chi_n(m_3ka_{23})\chi_n'(m_2ka_{23}) - \delta\,\chi_n'(m_3ka_{23})\chi_n(m_2ka_{23}), \tag{7.2d}$$

where $\gamma = m_2$ and $\delta = m_3$ for TE scattering, while $\gamma = m_3$ and $\delta = m_2$ for TM scattering. The four fundamental amplitudes for each interface satisfy the Wronskian relations

$$N_n^{12}Q_n^{12} - D_n^{12}P_n^{12} = m_1m_2 \tag{7.3a}$$

$$N_n^{23}Q_n^{23} - D_n^{23}P_n^{23} = m_2m_3. \tag{7.3b}$$

One can also define four fundamental partial wave amplitudes for the composite particle with the 12 and 23 interfaces lumped together as a single unit

$$N_n^{123} = D_n^{12}N_n^{23} - N_n^{12}P_n^{23} \tag{7.4a}$$

$$D_n^{123} = D_n^{12}D_n^{23} - N_n^{12}Q_n^{23} \tag{7.4b}$$

$$P_n^{123} = Q_n^{12}N_n^{23} - P_n^{12}P_n^{23} \tag{7.4c}$$

$$Q_n^{123} = Q_n^{12}D_n^{23} - P_n^{12}Q_n^{23}. \tag{7.4d}$$

These composite amplitudes satisfy the Wronskian relation

$$N_n^{123} Q_n^{123} - D_n^{123} P_n^{123} = m_1 m_2^2 m_3, \tag{7.5}$$

and will be called multiple-interface amplitudes. The compact form of the partial wave scattering amplitudes for the coated sphere may then be written as

$$a_n, b_n = N_n^{123} / \left(N_n^{123} + \mathrm{i}\, D_n^{123} \right). \tag{7.6}$$

The expressions for the partial wave amplitudes in the core and coating are similar. It should be noted that Eq. (7.6) has the same formal structure as Eq. (2.18) for a homogeneous sphere.

In order to set the stage for the Debye series expansion, one must first determine the partial wave reflection and transmission amplitudes at each of the interfaces. In analogy to Eqs. (3.4a, 3.4b, 3.7a, 3.7b) the derivation gives

$$T_n^{21} = (m_1/m_2) T_n^{12} = -2\mathrm{i}m_1 / \left[\left(N_n^{12} + Q_n^{12} \right) + \mathrm{i} \left(D_n^{12} - P_n^{12} \right) \right] \tag{7.7a}$$

$$R_n^{212} = \left[\left(-N_n^{12} + Q_n^{12} \right) + \mathrm{i} \left(D_n^{12} + P_n^{12} \right) \right] / \left[\left(N_n^{12} + Q_n^{12} \right) + \mathrm{i} \left(D_n^{12} - P_n^{12} \right) \right] \tag{7.7b}$$

$$R_n^{121} = \left[\left(-N_n^{12} + Q_n^{12} \right) - \mathrm{i} \left(D_n^{12} + P_n^{12} \right) \right] / \left[\left(N_n^{12} + Q_n^{12} \right) + \mathrm{i} \left(D_n^{12} - P_n^{12} \right) \right] \tag{7.7c}$$

for the 12 interface and

$$T_n^{32} = (m_2/m_3) T_n^{23} = -2\mathrm{i}m_2 / \left[\left(N_n^{23} + Q_n^{23} \right) + \mathrm{i} \left(D_n^{23} - P_n^{23} \right) \right] \tag{7.8a}$$

$$R_n^{323} = \left[\left(-N_n^{23} + Q_n^{23} \right) + \mathrm{i} \left(D_n^{23} + P_n^{23} \right) \right] / \left[\left(N_n^{23} + Q_n^{23} \right) + \mathrm{i} \left(D_n^{23} - P_n^{23} \right) \right] \tag{7.8b}$$

$$R_n^{232} = \left[\left(-N_n^{23} + Q_n^{23} \right) - \mathrm{i} \left(D_n^{23} + P_n^{23} \right) \right] / \left[\left(N_n^{23} + Q_n^{23} \right) + \mathrm{i} \left(D_n^{23} - P_n^{23} \right) \right] \tag{7.8c}$$

for the 23 interface. These partial wave transmission and reflection amplitudes will hereafter be called the single-interface amplitudes.

There are at least three different but equivalent ways to express the Debye series expansion of the partial wave scattering amplitudes of Eq. (7.6). First, the Debye series can be written in expanded form entirely in terms of single-interface amplitudes. Using the derivation outlined in Sect. 4 and again reading the ordering of the R and T factors from left to right as the partial wave progresses through the coated sphere, one obtains

$$\begin{aligned} a_n, b_n &= (1/2) \left[1 - R_n^{323} - T_n^{32} W_n^{212} \left(1 - R_n^{232} W_n^{212} \right)^{-1} T_n^{23} \right] \\ &= (1/2) \left[1 - R_n^{323} - \sum_{q=1}^{\infty} T_n^{32} W_n^{212} \left(R_n^{232} W_n^{212} \right)^{q-1} T_n^{23} \right], \end{aligned} \tag{7.9}$$

where

$$\begin{aligned} W_n^{212} &= R_n^{212} + T_n^{21}\left(1-R_n^{121}\right)^{-1}T_n^{12} \\ &= R_n^{212} + \sum_{p=1}^{\infty} T_n^{21}\left(R_n^{121}\right)^{p-1}T_n^{12}. \end{aligned} \tag{7.10}$$

The interpretation of the terms in Eq. (7.9) is that an incident partial wave can be diffracted by the coated sphere or externally reflected from its outer surface ($1-R_n^{232}$). It can also be transmitted through the outer surface (T_n^{32}), interact with the core (W_n^{212}) and then be transmitted back out (T_n^{23}). Alternatively, after it has interacted at the 12 interface it can propagate in region 2 alternately outward and inward $q-1$ times, reflecting from the 23 interface (R_n^{232}) and then interacting with the core (W_n^{212}), before being transmitted back out (T_n^{23}). When the partial wave in region 2 interacts with the core (W_n^{212}) in Eq. (7.10), it is directly reflected from the interface (R_n^{212}), or it is transmitted into region 1 (T_n^{21}), makes $p-1$ internal reflections within the core (R_n^{121}), and is then transmitted back into region 2 (T_n^{12}). These interactions are pictorially illustrated in Fig. 6.

Second, the Debye series can be expressed in compact form in terms of multiple-scattering amplitudes as

$$\begin{aligned} a_n, b_n &= (1/2)[1-\mathcal{R}_n^{3\gamma 3}-\mathcal{T}_n^{31}(1-\mathcal{R}_n^{1\gamma 1})^{-1}\mathcal{T}_n^{13}] \\ &= (1/2)[1-\mathcal{R}_n^{3\gamma 3} - \sum_{p=1}^{\infty}\mathcal{T}_n^{31}(\mathcal{R}_n^{1\gamma 1})^{p-1}\mathcal{T}_n^{13}], \end{aligned} \tag{7.11}$$

where the multiple-scattering partial wave transmission and reflection amplitudes in Eq. (7.11) are defined in analogy to Eqs. (7.7a–7.7c, 7.8a–7.8c) as

$$\begin{aligned} \mathcal{T}_n^{31} &= 2im_1m_2/\left[\left(N_n^{123}+Q_n^{123}\right)+\mathrm{i}\left(D_n^{123}-P_n^{123}\right)\right] \\ &= T_n^{32}\left(1-R_n^{212}R_n^{232}\right)^{-1}T_n^{21} \end{aligned} \tag{7.12a}$$

$$\begin{aligned} \mathcal{T}_n^{13} &= 2im_2m_3/\left[\left(N_n^{123}+Q_n^{123}\right)+\mathrm{i}\left(D_n^{123}-P_n^{123}\right)\right] \\ &= T_n^{12}\left(1-R_n^{232}R_n^{212}\right)^{-1}T_n^{23} \end{aligned} \tag{7.12b}$$

$$\begin{aligned} \mathcal{R}_n^{3\gamma 3} &= \left[\left(-N_n^{123}+Q_n^{123}\right)+\mathrm{i}\left(D_n^{123}+P_n^{123}\right)\right]/\left[\left(N_n^{123}+Q_n^{123}\right)+\mathrm{i}\left(D_n^{123}-P_n^{123}\right)\right] \\ &= R_n^{323}+T_n^{32}R_n^{212}\left(1-R_n^{232}R_n^{212}\right)^{-1}T_n^{23} \end{aligned} \tag{7.12c}$$

$$\begin{aligned} \mathcal{R}_n^{1\gamma 1} &= \left[\left(-N_n^{123}+Q_n^{123}\right)-\mathrm{i}\left(D_n^{123}+P_n^{123}\right)\right]/\left[\left(N_n^{123}+Q_n^{123}\right)+\mathrm{i}\left(D_n^{123}-P_n^{123}\right)\right] \\ &= R_n^{121}+T_n^{12}R_n^{232}\left(1-R_n^{212}R_n^{232}\right)^{-1}T_n^{21}. \end{aligned} \tag{7.12d}$$

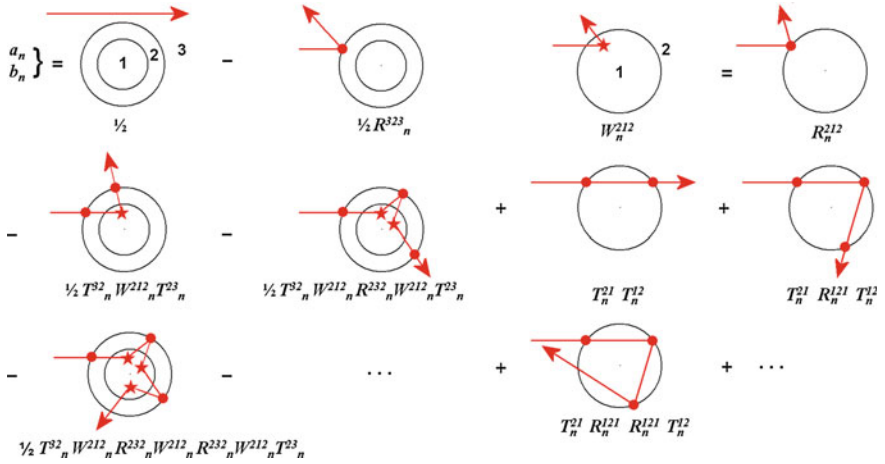


Fig. 6 Terms of the Debye series expansion of the partial wave scattering amplitudes a_n, b_n for a coated sphere, and of the core interaction amplitude W_n^{212}

The multiple-scattering transmission amplitudes $\mathcal{T}_n^{31}$ and $\mathcal{T}_n^{13}$ correspond to a partial wave being transmitted from either region 3 or region 1 into region 2 (T_n^{32} or T_n^{12}), then reflecting back and forth in region 2 any number of times ($R_n^{212}R_n^{232}$ or $R_n^{232}R_n^{212}$), and finally being transmitted into region 1 or region 3 (T_n^{21} or T_n^{23}). Similarly, the multiple-scattering reflection amplitudes $\mathcal{R}_n^{3\gamma 3}$ and $\mathcal{R}_n^{1\gamma 1}$ correspond to a partial wave in either region 3 or region 1 being directly reflected back (R_n^{323} or R_n^{121}) or transmitted into region 2 and reflecting from the 12 or 23 interface ($T_n^{32}R_n^{121}$ or $T_n^{12}R_n^{232}$), then reflecting back and forth in region 2 any number of times ($R_n^{232}R_n^{212}$ or $R_n^{212}R_n^{232}$), and finally transmitted back into region 3 or region 1 (T_n^{23} or T_n^{21}). The compact form Debye series of Eq. (7.11) has the same formal structure as for a homogeneous sphere in Eq. (2.18), except that it is written in terms of multiple-scattering amplitudes rather than single-interface amplitudes.

As an example of the use of the coated sphere Debye series, the first-order rainbow becomes a closely spaced pair of twin first-order rainbows. The α component

$$a_n, b_n = (-1/2)T_n^{32}T_n^{21}R_n^{121}T_n^{12}T_n^{23} \tag{7.13}$$

has its internal reflection at the 12 interface, and the β component

$$a_n, b_n = (-1/2)T_n^{32}T_n^{21}T_n^{12}R_n^{232}T_n^{21}T_n^{12}T_n^{23} \tag{7.14}$$

has its internal reflection at the 23 interface. The evolution of these twin first-order rainbows was examined in detail as a function of coating thickness in Lock et al. (1994). Similarly, there are three closely spaced components of the second-order

rainbow (Laven and Lock 2012), where (i) both internal reflections occur at the 12 interface, $\alpha\alpha$, (ii) both at the 23 interface, $\beta\beta$, and (iii) one each at the 12 and 23 interfaces. For this last possibility, the first of the internal reflections is either at the 12 interface and the second is at the 23 interface, $\alpha\beta$, or the first is at the 23 interface and the second is at the 12 interface, $\beta\alpha$. Each of these two orderings gives an identical contribution to a_n and b_n.

This multiplicity leads to the third way of expressing the Debye series that avoids having to evaluate a large number of terms that give exactly the same contribution to a_n and b_n. Let N be the total number of internal reflections for a given term of the Debye series, N^{212} the number of internal reflections at the 12 interface incident from region 2, N^{121} the number at the 12 interface incident from region 1, and N^{232} the number at the 23 interface incident from region 2, with

$$N = N^{212} + N^{121} + N^{232}. \tag{7.15}$$

We now reorganize the expanded form of the Debye series from specifying the values of p and q of Eqs. (7.9, 7.10) to specifying the total number of internal reflections N. For $0 \le N \le 7$, there are 2, 3, 6, 14, 31, 70, 157, 353 different Debye series terms. For example, for $N = 0$ there is

$$a_n, b_n = (1/2)\left[\left(1 - R_n^{323}\right) - T_n^{32} T_n^{21} T_n^{12} T_n^{23}\right] \tag{7.16}$$

i.e., diffraction-plus-external reflection, and direct transmission in and out. For $N = 1$ there are the two twin first-order rainbow terms of Eqs. (7.13,7.14), as well as reflection at the 12 interface without the partial wave having ever entered the core,

$$a_n, b_n = (-1/2) T_n^{32} R_n^{212} T_n^{23}. \tag{7.17}$$

The 25 terms for $0 \le N \le 3$ are pictorially illustrated in Fig. 7. Having to calculate all these terms for even a moderate value of N quickly becomes extremely inefficient.

In addition to categorizing the individual terms of the Debye series terms by their value of N, they are also categorized by the number of chords A the corresponding ray path has in the coating, and the number of chords B the ray path has in the core. One finds that

$$N^{232} = (A - 2)/2 \tag{7.18a}$$

$$N^{121} = (N + 1 - A + B)/2 \tag{7.18b}$$

$$N^{212} = (N + 1 - B)/2. \tag{7.18c}$$

Each Debye series term with the same value of N, A, B is merely a different reordering of all the partial wave transmission and reflection amplitudes involved,

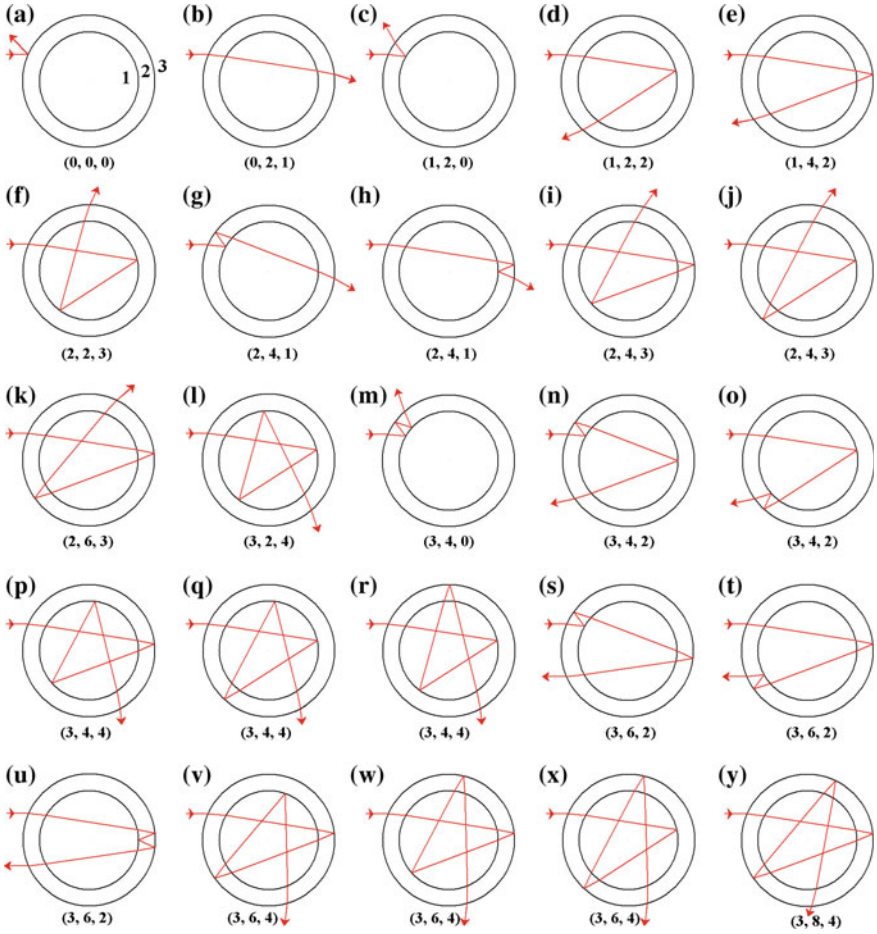


Fig. 7 $0 \le N \le 3$ terms of the Debye series expansion of the partial wave scattering amplitudes a_n, b_n for a coated sphere, showing their (N, A, B) values where A is the number of chords in the coating and B is the number of chords in the core

and gives the same contribution to a_n and b_n. Counting in this way and taking into account the degeneracy factor $D(N, N^{232}, N^{212})$ of the different but equivalent contributions to a_n and b_n, the number of distinct terms for $0 \le N \le 7$ now is reduced to 2, 3, 4, 7, 9, 13, 16, 21, which is a much more manageable computational task. The reorganized expanded form of the Debye series then becomes

$$a_n, b_n = (1/2)\left(1 - R_n^{323} - \sum_{N=0}^{\infty} T_n^{32} F_n^N T_n^{23}]\right), \tag{7.19}$$

where

$$F_n^N = \sum_{N^{232}=0}^{N} \sum_{N^{212}=0}^{N_{\max}} D\left(N, N^{232}, N^{212}\right)\left(T_n^{21} T_n^{12}\right)^{N232+1-N212}\left(R_n^{232}\right)^{N232}\left(R_n^{212}\right)^{N212}\left(R_n^{121}\right)^{N121} + G_n, \tag{7.20}$$

$N_{\max}$ is the smaller of N^{232} and $N - N^{232}$,

$$\begin{aligned} G_n &= \left(R_n^{232}\right)^{(N-1)/2}\left(R_n^{212}\right)^{(N+1)/2} && \text{for odd } N \\ &= 0 && \text{for even } N, \end{aligned} \tag{7.21}$$

and the ray path degeneracy factor is

$$D = \left(N^{232}+1\right)!\left(N-2N^{212}\right)!/\left[\left(N^{212}\right)!\left(N^{232}+1-N^{212}\right)!\left(N-N^{232}-N^{212}\right)!\left(N^{232}-N^{212}\right)!\right]. \tag{7.22}$$

In Lock and Laven (2012), the scattered intensity corresponding to Eqs. (7.19–7.22) was computed for $0 \le N \le 3$ and compared with that of the exact solution of Eq. (7.6). The agreement was found to be nearly perfect.

7.2 Scattering by a Multi-layer Sphere

We now turn our attention to scattering of an electromagnetic plane wave by a stratified sphere containing $M \ge 3$ concentric layers, where region $M + 1$ is the exterior medium. The partial wave scattering amplitudes again are found to take the form

$$a_n, b_n = N_n^{12\ldots M+1}/\left(N_n^{12\ldots M+1} + \mathrm{i}\, D_n^{12\ldots M+1}\right), \tag{7.23}$$

in analogy to Eq. (2.18) for a homogeneous sphere and Eq. (7.6) for a singly coated sphere. The four fundamental multiple-interface amplitudes $N_n^{12\ldots M+1}$, $D_n^{12\ldots M+1}$, $P_n^{12\ldots M+1}$, $Q_n^{12\ldots M+1}$ in Eq. (7.23) can be obtained by progressive iteration, starting from the innermost layer and working outward, adding one interface at a time. This is done by using the multiple-interface amplitudes for the combination starting from interface 12 to interface $J - 1$, J, called $12\ldots J$ for short, along with single-interface amplitudes for the J, $J + 1$ interface,

$$N_n^{12\ldots J+1} = D_n^{12\ldots J} N_n^{J,J+1} - N_n^{12\ldots J} P_n^{J,J+1} \tag{7.24a}$$

$$D_n^{12\ldots J+1} = D_n^{12\ldots J} D_n^{J,J+1} - N_n^{12\ldots J} Q_n^{J,J+1} \tag{7.24b}$$

$$P_n^{12\ldots J+1} = Q_n^{12\ldots J} N_n^{J,J+1} - P_n^{12\ldots J} P_n^{J,J+1} \tag{7.24c}$$

$$Q_n^{12\ldots J+1} = Q_n^{12\ldots J} D_n^{J,J+1} - P_n^{12\ldots J} Q_n^{J,J+1}, \tag{7.24d}$$

in analogy to Eqs. (7.4a–7.4d) for a singly coated sphere. This is a variant of the progressive iteration scheme for a multi-layer sphere or a stack of many plane parallel layers [see pp. 51–70 of Born and Wolf (1980), pp. 45–53 of Chew (1990), Bhandari (1985), Mackowski et al. (1990), Wu and Wang (1991), Kai and Massoli (1994), Johnson (1996), and Wu et al. (1997)].

The Debye series of Eq. (7.23) for a multi-layer sphere can be made to resemble either the expanded form of Eqs. (7.9, 7.10) or the compact form of Eq. (7.11) for a singly coated sphere. For the expanded form, one uses multiple-scattering transmission and reflection amplitudes for the first M interfaces $\mathcal{T}_n^{M,1}$, $\mathcal{T}_n^{1,M}$, $\mathcal{R}_n^{1,\gamma,1}$, and $\mathcal{R}_n^{M,\gamma,M}$, and the single-interface transmission and reflection amplitudes at the outer surface of the sphere, i.e., the M, $M+1$ interface. The multiple-scattering transmission and reflection amplitudes are identical to Eqs. (3.4a, 3.4b, 3.7a, 3.7b) for a homogeneous sphere and Eqs. (7.12a–7.12d) for a singly coated sphere, except they use $N_n^{12\ldots M}$, $D_n^{12\ldots M}$, $P_n^{12\ldots M}$, $Q_n^{12\ldots M}$ rather than N_n, D_n, P_n, Q_n or N_n^{123}, D_n^{123}, P_n^{123}, Q_n^{123}. The symbol γ denotes multiple-scattering within the first $M-1$ layers, consisting of all the combinations of the interactions at all the intervening interfaces. Alternatively, the multiple-scattering amplitudes can be obtained by progressive iteration

$$\mathcal{T}_n^{M,1} = T_n^{M,M-1}\left(1 - \mathcal{R}_n^{M-1,\delta,M-1} R_n^{M-1,M,M-1}\right)^{-1} \mathcal{T}_n^{M-1,1} \tag{7.25a}$$

$$\mathcal{T}_n^{1,M} = \mathcal{T}_n^{1,M-1}\left(1 - \mathcal{R}_n^{M-1,\delta,M-1} R_n^{M-1,M,M-1}\right)^{-1} T_n^{M-1,M} \tag{7.25b}$$

$$\mathcal{R}_n^{M,\gamma,M} = R_n^{M,M-1,M} + T_n^{M,M-1}\mathcal{R}_n^{M-1,\delta,M-1}\left(1 - \mathcal{R}_n^{M-1,\delta,M-1} R_n^{M-1,M,M-1}\right)^{-1} T_n^{M-1,M} \tag{7.25c}$$

$$\mathcal{R}_n^{1,\gamma,1} = \mathcal{R}_n^{1,\delta,1} + \mathcal{T}_n^{1,M-1} R_n^{M-1,M,M-1}\left(1 - \mathcal{R}_n^{M-1,\delta,M-1} R_n^{M-1,M,M-1}\right)^{-1} \mathcal{T}_n^{M-1,1}, \tag{7.25d}$$

in generalization of the second part of Eqs. (7.12a–7.12d) for a singly coated sphere, where δ denotes multiple-scattering within the first $M-2$ layers. The Debye series then becomes

$$\begin{aligned} a_n, b_n &= (1/2)\left[1 - R_n^{M+1,M,M+1} - T_n^{M+1,M}\mathcal{W}_n^M\left(1 - R_n^{M,M+1,M}\mathcal{W}_n^M\right)^{-1} T_n^{M,M+1}\right] \\ &= (1/2)\left[1 - R_n^{M+1,M,M+1} - \sum_{q=1}^{\infty} T_n^{M+1,M}\mathcal{W}_n^M\left(R_n^{M,M+1,M}\mathcal{W}_n^M\right)^{q-1} T_n^{M,M+1}\right] \end{aligned} \tag{7.26}$$

where the multiple scattering amplitude $\mathcal{W}_n^M$ is given by

$$\begin{aligned}\mathcal{W}_n^M &= \mathcal{R}_n^{M,\gamma,M} + \mathcal{T}_n^{M,1}\left(1-\mathcal{R}_n^{1,\gamma,1}\right)^{-1}\mathcal{T}_n^{1,M} \\ &= \mathcal{R}_n^{M,\gamma,M} + \sum_{p=1}^{\infty}\mathcal{T}_n^{M,1}\left(\mathcal{R}_n^{1,\gamma,1}\right)^{p-1}\mathcal{T}_n^{1,M}.\end{aligned} \tag{7.27}$$

Equation (7.27) for $\mathcal{W}_n^M$ can be interpreted as the partial wave interacting with the first $M-1$ layers of the stratified sphere in analogy to (7.10) for a singly coated sphere, with the diffraction term subtracted out. Equation (7.26) describes the effect of adding on the final layer, with $T_n^{M+1,M}$, $T_n^{M,M+1}$, $R_n^{M+1,M,M+1}$, and $R_n^{M,M+1,M}$ being single-interface partial wave transmission and reflection amplitudes. In this way, individual layers can be added on one by one in order to form the multi-layer sphere.

For the compact form of the Debye series one uses multiple-scattering transmission and reflection amplitudes for all M interfaces from the core out to the surface of the sphere, here denoted by Γ. Using an analysis similar to that for the situation discussed in the previous paragraph, the result turns out to be

$$\begin{aligned}a_n, b_n &= (1/2)\left[1-\mathcal{R}_n^{M+1,\Gamma,M+1}-\mathcal{T}_n^{M+1,1}\left(1-\mathcal{R}_n^{1,\Gamma,1}\right)^{-1}\mathcal{T}_n^{1,M+1}\right] \\ &= (1/2)\left[1-\mathcal{R}_n^{M+1,\Gamma,M+1} - \sum_{p=1}^{\infty}\mathcal{T}_n^{M+1,1}\left(\mathcal{R}_n^{1,\Gamma,1}\right)^{p-1}\mathcal{T}_n^{1,M+1}\right]\end{aligned} \tag{7.28}$$

It is both physically and mathematically pleasing that the patterns observed in the formulas for the partial wave scattering amplitudes for a homogeneous sphere and a singly coated sphere continue in exactly the same way for scattering by a multi-layer sphere when single-interface amplitudes are replaced by the corresponding multiple-scattering amplitudes. The same patterns will be seen to continue when polarization-preserving and polarization-changing amplitudes are included for the geometry of Sect. 8, and when partial wave coupling is included for the geometry of Sect. 9.

8 Debye Series for a Plane Wave Diagonally Incident on a Cylinder

Consider a plane wave in the external medium 2 incident on an infinitely long circular cylinder of radius a and refractive index m_1 whose symmetry axis coincides with the z axis of a rectangular coordinate system. The scattering angle θ lies in the xy plane and is measured with respect to the $+x$ axis. If the plane wave is normally incident on the cylinder with its propagation direction along the x axis, the scattered fields are relatively simple. The partial wave scattering amplitudes a_n and b_n have

exactly the same form as for scattering by a sphere, but with the Bessel, Neumann, and Hankel functions J_n, N_n, $H_n^{(1)}$, and $H_n^{(2)}$ replacing the spherical Bessel, Neumann, and Hankel functions j_n, n_n, $h_n^{(1)}$, and $h_n^{(2)}$ [see Sect. 15.12 of van de Hulst (1981), and Sect. 6.1.3 of Kerker (1969)]. But if the plane wave propagates in the xz plane with the wave vector

$$\mathbf{k} = k[\cos(\zeta)\mathbf{u_x} - \sin(\zeta)\mathbf{u_z}], \tag{8.1}$$

and is diagonally incident on the cylinder, the situation becomes much more complicated because polarization-preserving and polarization-changing scattering now occur. To describe this situation, we first need to specify a pair of convenient polarization states of the incident beam and the scattered radiation [see pp. 298–299 of van de Hulst (1981) and pp. 256–257 of Kerker (1969)]. The ε polarization has the electric field vector in the xy plane, while the μ polarization has the magnetic field vector in the xy plane. For the incident beam described in terms of rectangular coordinates this means that

$$\mathbf{E^{inc,\varepsilon}} = E_0\,\mathbf{u_y}\exp[ikx\,\cos(\zeta) - ikz\,\sin(\zeta)] \tag{8.2a}$$

$$(c/m_2)\mathbf{B^{inc,\varepsilon}} = E_0[\sin(\zeta)\mathbf{u_x} + \cos(\zeta)\mathbf{u_z}]\exp[ikx\,\cos(\zeta) - ikz\,\sin(\zeta)] \tag{8.2b}$$

and

$$\mathbf{E^{inc,\mu}} = E_0[\sin(\zeta)\mathbf{u_x} + \cos(\zeta)\mathbf{u_z}]\exp[ikx\cos(\zeta) - ikz\sin(\zeta)] \tag{8.3a}$$

$$(c/m_2)\mathbf{B^{inc,\mu}} = -E_0\mathbf{u_y}\exp[ikx\cos(\zeta) - ikz\sin(\zeta)]. \tag{8.3b}$$

It is more convenient, however, to express the far-zone scattered fields in cylindrical coordinates ρ, θ, z. The polarization-preserving $\varepsilon \rightarrow \varepsilon$ scattered fields are

$$\mathbf{E^{scatt,\varepsilon\varepsilon}} = S_{\varepsilon\varepsilon}(\rho,\theta,z)\mathbf{u_\theta} \tag{8.4a}$$

$$(c/m_2)\mathbf{B^{scatt,\varepsilon\varepsilon}} = S_{\varepsilon\varepsilon}(\rho,\theta,z)\left[\sin(\zeta)\mathbf{u_\rho} + \cos(\zeta)\mathbf{u_z}\right], \tag{8.4b}$$

where $S_{\varepsilon\varepsilon}$ is the $\varepsilon \rightarrow \varepsilon$ scattering amplitude to be obtained below. The polarization subscripts here, as well as the polarization superscripts in Eqs. (8.17a–8.17c to 8.23a–8.23c, 8.23d), read from left to right, i.e., ε-incoming goes to ε-outgoing. The polarization-preserving $\mu \rightarrow \mu$ scattered fields are

$$\mathbf{E^{scatt,\mu\mu}} = S_{\mu\mu}(\rho,\theta,z)\left[\sin(\zeta)\mathbf{u_\rho} + \cos(\zeta)\mathbf{u_z}\right] \tag{8.5a}$$

$$(c/m_2)\mathbf{B^{scatt,\mu\mu}} = S_{\mu\mu}(\rho,\theta,z)(-\mathbf{u_\theta}) \tag{8.5b}$$

where $S_{\mu\mu}$ is the $\mu \rightarrow \mu$ scattering amplitude. It will turn out that the polarization-changing $\varepsilon \rightarrow \mu$ and $\mu \rightarrow \varepsilon$ scattering amplitudes S_q are opposite in sign, so that

$$\mathbf{E}^{\mathbf{scatt},\varepsilon\mu} = S_q(\rho,\theta,z)\left[\sin(\zeta)\mathbf{u}_{\boldsymbol{\rho}} + \cos(\zeta)\mathbf{u}_{\mathbf{z}}\right] \tag{8.6a}$$

$$(c/m_2)\mathbf{B}^{\mathbf{scatt},\varepsilon\mu} = S_q(\rho,\theta,z)(-\mathbf{u}_{\boldsymbol{\theta}}) \tag{8.6b}$$

and

$$\mathbf{E}^{\mathbf{scatt},\mu\varepsilon} = -S_q(\rho,\theta,z)\mathbf{u}_{\boldsymbol{\theta}} \tag{8.7a}$$

$$(c/m_2)\mathbf{B}^{\mathbf{scatt},\mu\varepsilon} = -S_q(\rho,\theta,z)\left[\sin(\zeta)\mathbf{u}_{\boldsymbol{\rho}} + \cos(\zeta)\mathbf{u}_{\mathbf{z}}\right]. \tag{8.7b}$$

Next we decompose the incident and scattered fields expressed in cylindrical coordinates as a sum over partial waves. The ε-polarized and μ-polarized scalar radiation potential is written as $\psi_n^X\left[h; m_j k(1-h^2)^{1/2}\rho, \theta, z\right]$ in cylindrical coordinates, where n is the integer partial wave number with $-\infty < n < \infty$, h is the continuous separation variable with $-\infty < h < \infty$ and $X = \varepsilon, \mu$. Since the function ψ^X satisfies the scalar wave equation, it can be written as a Bessel, Neumann, or Hankel function of $m_j k(1-h^2)^{1/2}\rho$ multiplied by the complex exponentials $\exp(in\theta)$ and $\exp(im_j khz)$, and an arbitrary complex constant giving the magnitude and phase of the partial wave contribution. The result is then summed over n and integrated over h for a general incident beam. However, only a single value of h corresponding to the tilt angle ζ occurs for an incident plane wave. The ε-polarized electric and magnetic fields associated with ψ^ε are given by

$$\mathbf{E}^{\boldsymbol{\varepsilon}} = -\mathbf{u}_{\mathbf{z}} \times \nabla\psi^{\varepsilon} \tag{8.8a}$$

$$\left(\mathrm{c}/m_j\right)\mathbf{B}^{\boldsymbol{\varepsilon}} = \left(-\mathrm{i}/m_j k\right)\nabla \times \mathbf{E}^{\boldsymbol{\varepsilon}}, \tag{8.8b}$$

and the μ-polarized electric and magnetic fields associated with ψ^μ are

$$\left(c/m_j\right)\mathbf{B}^{\boldsymbol{\mu}} = \mathbf{u}_{\mathbf{z}} \times \nabla\psi^{\mu} \tag{8.9a}$$

$$\mathbf{E}^{\boldsymbol{\mu}} = \left[\mathrm{i}c/\left(m_j^2 k\right)\right]\nabla \times \mathbf{B}^{\boldsymbol{\mu}}. \tag{8.9b}$$

Again Eqs. (8.8a, 8.9a) are frequently associated with the vector cylindrical wave function $\mathbf{M}_{\mathbf{n}}(h; k\mathbf{r})$, and Eqs. (8.8b, 8.9b) are associated with $\mathbf{N}_{\mathbf{n}}(h; k\mathbf{r})$.

In the far-zone, the polarization-preserving and crossed-polarized scattering amplitudes in Eqs. (8.4a, 8.4b–8.7a, 8.7b) separate into the product form

$$S_{\varepsilon\varepsilon}(\rho,\theta,z) = S^{\mathrm{PW}}(\rho,z)\left[a_0 + 2\sum_{n=1}^{\infty} a_n \cos(n\theta)\right] \tag{8.10a}$$

$$S_{\mu\mu}(\rho,\theta,z) = S^{\mathrm{PW}}(\rho,z)\left[b_0 + 2\sum_{n=1}^{\infty} b_n \cos(n\theta)\right] \tag{8.10b}$$

$$S_q(\rho,\theta,z) = S^{\mathrm{PW}}(\rho,z)\left[2\mathrm{i}\sum_{n=1}^{\infty} q_n \sin(n\theta)\right] \tag{8.10c}$$

where a_n, b_n, q_n, are the partial wave scattering amplitudes and

$$S^{\mathrm{PW}}(\rho,z) = -E_0\{2/[\pi k\,\rho\cos(\zeta)]\}^{1/2}\exp(-\mathrm{i}\pi/4)\exp[\mathrm{i}k\,\rho\cos(\zeta) - \mathrm{i}kz\sin(\zeta)]. \tag{8.11}$$

The calculation of the partial wave scattering amplitudes for a plane wave diagonally incident on an infinitely long circular cylinder is complicated by the fact that the symmetry that occurred for normal beam incidence is now broken (Wait 1955). Both the partial wave scattering amplitudes and the Debye series terms can be conveniently written in terms of the fundamental amplitudes,

$$E_n^{(i,j)} = m_2 ka\cos(\zeta)\cos(\chi)\left[m_2\cos(\zeta)H_n^{(i)}(x)H_n^{(j)\prime}(y) - m_1\cos(\chi)H_n^{(i)\prime}(x)H_n^{(j)}(y)\right] \tag{8.12a}$$

$$M_n^{(i,j)} = m_1 ka\cos(\zeta)\cos(\chi)\left[m_1\cos(\zeta)H_n^{(i)}(x)H_n^{(j)\prime}(y) - m_2\cos(\chi)H_n^{(i)\prime}(x)H_n^{(j)}(y)\right] \tag{8.12b}$$

$$Q_n^{(i,j)} = n\left[m_1\sin(\zeta)\cos^2(\chi) - m_2\sin(\chi)\cos^2(\zeta)\right]H_n^{(i)}(x)H_n^{(j)}(y) \tag{8.12c}$$

and

$$D_n^{(i,j)} = E_n^{(i,j)}M_n^{(i,j)} - \left(Q_n^{(i,j)}\right)^2. \tag{8.12d}$$

In Eqs. (8.12a–8.12d) one has $i = 0, 1, 2$ and $j = 0, 1, 2$ with the notational shorthand $H_n^{(0)}(x) = J_n(x)$ and $H_n^{(0)}(y) = J_n(y)$. The cylinder size parameters are

$$x = m_2 ka\,\cos(\zeta). \tag{8.13a}$$

$$y = m_1 ka\,\cos(\chi), \tag{8.13b}$$

and the angles ζ and χ are related by Snell's law

$$m_2\,\sin(\zeta) = m_1\,\sin(\chi). \tag{8.14}$$

The partial wave polarization-preserving and cross-polarized scattering amplitudes are then found to be

$$a_n = \left(E_n^{(0,0)} M_n^{(1,0)} - Q_n^{(0,0)} Q_n^{(1,0)}\right) / D_n^{(1,0)} \tag{8.15a}$$

$$b_n = \left(M_n^{(0,0)} E_n^{(1,0)} - Q_n^{(0,0)} Q_n^{(1,0)}\right) / D_n^{(1,0)} \tag{8.15b}$$

$$q_n = \left[2m_1 \cos^2(\chi)/\pi\right] Q_n^{(0,0)} J_n(y) / \left[D_n^{(1,0)} J_n(x)\right], \tag{8.15c}$$

As an aside, for normal plane wave incidence one has $\sin(\zeta) = \sin(\chi) = 0$ and $\cos(\zeta) = \cos(\chi) = 1$. Then $a_n = E_n^{(0,0)}/E_n^{(1,0)}$ for $\varepsilon \rightarrow \varepsilon$ polarization-preserving scattering and $b_n = M_n^{(0,0)}/M_n^{(1,0)}$ for $\mu \rightarrow \mu$ polarization-preserving scattering. In addition $Q_n^{(0,0)} = Q_n^{(1,0)} = Q_n^{(i,j)} = 0$, which gives $q_n = 0$ and forbids cross-polarized $\varepsilon \rightarrow \mu$ and $\mu \rightarrow \varepsilon$ scattering at normal incidence. It should also be noted that for scattering of a scalar plane wave by a *sphere*, the partial wave scattering amplitudes were equal to the TE version for scattering by an electromagnetic plane wave. Similarly, for scattering of a normally incident scalar plane wave by a *cylinder*, the partial wave scattering amplitudes were equal to the μ-polarized version for scattering of a normally incident electromagnetic wave. This pattern, however, does not extend to diagonal incidence on a cylinder where the scalar wave scattering amplitudes are

$$\begin{aligned} b_n = &\left[m_1 \cos(\chi) J_n(x) J_n'(y) - m_2 \cos(\zeta) J_n'(x) J_n(y)\right] \\ &/ \left[m_1 \cos(\chi) H_n^{(1)}(x) J_n'(y) - m_2 \cos(\zeta) H_n^{(1)\prime}(x) J_n(y)\right]. \end{aligned} \tag{8.16}$$

This quantity is not equal to $M_n^{(0,0)}/M_n^{(1,0)}$. The difference is attributable to the different boundary conditions that scalar waves and electromagnetic waves are subject to.

As was the case in Sect. 3, the first step in the derivation of the Debye series of a_n, b_n, q_n is to determine the polarization-preserving and polarization-changing transmission and reflection amplitudes when a single cylindrically incoming or outgoing ε-polarized or μ-polarized partial wave is incident on the cylinder. For the 212 exterior reflection, one obtains

$$R_n^{212,\varepsilon\varepsilon} = -\left(E_n^{(2,2)} M_n^{(1,2)} - Q_n^{(1,2)} Q_n^{(2,2)}\right) / D_n^{(1,2)} \tag{8.17a}$$

$$R_n^{212,\mu\mu} = -\left(E_n^{(1,2)} M_n^{(2,2)} - Q_n^{(1,2)} Q_n^{(2,2)}\right) / D_n^{(1,2)} \tag{8.17b}$$

$$R_n^{212,\varepsilon\mu} = -R_n^{212,\mu\varepsilon} = \left[4m_1 \cos^2(\chi)/\pi\right] Q_n^{(1,2)} H_n^{(2)}(y) / \left[D_n^{(1,2)} H_n^{(1)}(x)\right]. \tag{8.17c}$$

For the 121 interior reflection, one has

$$R_n^{121,\varepsilon\varepsilon} = -\left(E_n^{(1,1)}M_n^{(1,2)} - Q_n^{(1,1)}Q_n^{(1,2)}\right)/D_n^{(1,2)} \tag{8.18a}$$

$$R_n^{121,\mu\mu} = -\left(E_n^{(1,2)}M_n^{(1,1)} - Q_n^{(1,1)}Q_n^{(1,2)}\right)/D_n^{(1,2)} \tag{8.18b}$$

$$R_n^{121,\varepsilon\mu} = -R_n^{121,\mu\varepsilon} = \left[4m_2\cos^2(\zeta)/\pi\right]Q_n^{(1,2)}H_n^{(1)}(x)/\left[D_n^{(1,2)}H_n^{(2)}(y)\right]. \tag{8.18c}$$

For the 21 transmission from the exterior medium to the cylinder interior one has

$$(m_1/m_2)T_n^{21,\varepsilon\varepsilon} = [-4\mathrm{i}m_2\cos^2(\zeta)/\pi]M_n^{(1,2)}/D_n^{(1,2)} \tag{8.19a}$$

$$T_n^{21,\mu\mu} = [-4\mathrm{i}m_2\cos^2(\zeta)/\pi]E_n^{(1,2)}/D_n^{(1,2)} \tag{8.19b}$$

$$T_n^{21,\varepsilon\mu} = (-m_1/m_2)T_n^{21,\mu\varepsilon} = [4m_2\cos^2(\zeta)/\pi]Q_n^{(1,2)}/D_n^{(1,2)}. \tag{8.19c}$$

Lastly, for the 12 transmission from the cylinder interior to the exterior medium one has

$$T_n^{12,\varepsilon\varepsilon} = \left[-4\mathrm{i}m_1\cos^2(\chi)/\pi\right]M_n^{(1,2)}/D_n^{(1,2)} \tag{8.20a}$$

$$(m_2/m_1)T_n^{12,\mu\mu} = [-4\mathrm{i}m_1\cos^2(\chi)/\pi]E_n^{(1,2)}/D_n^{(1,2)} \tag{8.20b}$$

$$T_n^{12,\varepsilon\mu} = (-m_2/m_1)T_n^{12,\mu\varepsilon} = [4m_1\cos^2(\chi)/\pi]Q_n^{(1,2)}/D_n^{(1,2)}. \tag{8.20c}$$

For a sphere or a cylinder at normal incidence there were eight partial wave transmission and reflection amplitudes that coupled together in two groups of four. The four TE amplitudes coupled together separately and the four TM amplitudes coupled together separately. For a coated sphere there were 16 partial wave transmission and reflection amplitudes which again coupled together in two groups of eight, the TE amplitudes separately and the TM amplitudes separately. For diagonal incidence on a cylinder there are also 16 partial wave transmission and reflection amplitudes. Since the symmetry of the previous geometries has now been broken, all 16 polarization-preserving and polarization-changing amplitudes couple together in the Debye series. But we will see in the next paragraph that their organization in the decomposition of a_n, b_n, q_n remains faithful to the patterns already observed for the higher symmetry geometries.

We let $\sum_\nu$ and $\sum_\tau$ denote a sum over the polarization states $\nu = \varepsilon, \mu$, and $\tau = \varepsilon, \mu$, and define the multiple-scattering internal reflection amplitude $\mathcal{R}_n^{121,\delta\sigma}$ as

$$\mathcal{R}_n^{121,\delta\sigma} = R_n^{121,\delta\sigma} + \sum_\nu R_n^{121,\delta\nu}R_n^{121,\nu\sigma} + \sum_\nu\sum_\tau R_n^{121,\delta\nu}R_n^{121,\nu\tau}R_n^{121,\tau\sigma} + \cdots, \tag{8.21}$$

reading both the polarization superscripts and the successive factors of R from left to right. This multiple scattering-amplitude includes all numbers of successive internal reflections of the partial wave inside the cylinder. Both polarization-preserving interactions $\varepsilon \rightarrow \varepsilon$ and $\mu \rightarrow \mu$ and polarization-changing interactions $\varepsilon \rightarrow \mu$ and $\mu \rightarrow \varepsilon$ occur at every interaction, with the constraint that the polarization is δ at the beginning and σ at the end. This is pictorially illustrated in Fig. 8.The series of Eq. (8.21) can be analytically summed to give

$$\begin{aligned}\mathcal{R}_n^{121,\varepsilon\varepsilon} = &\left[R_n^{121,\varepsilon\varepsilon}\left(1 - R_n^{121,\mu\mu}\right) + R_n^{121,\varepsilon\mu}R_n^{121,\mu\varepsilon}\right] \\ &/\left[\left(1 - R_n^{121,\varepsilon\varepsilon}\right)\left(1 - R_n^{121,\mu\mu}\right) - R_n^{121,\varepsilon\mu}R_n^{121,\mu\varepsilon}\right]\end{aligned} \tag{8.22a}$$

$$\mathcal{R}_n^{121,\varepsilon\mu} = R_n^{121,\varepsilon\mu}/\left[\left(1 - R_n^{121,\varepsilon\varepsilon}\right)\left(1 - R_n^{121,\mu\mu}\right) - R_n^{121,\varepsilon\mu}R_n^{121,\mu\varepsilon}\right]. \tag{8.22b}$$

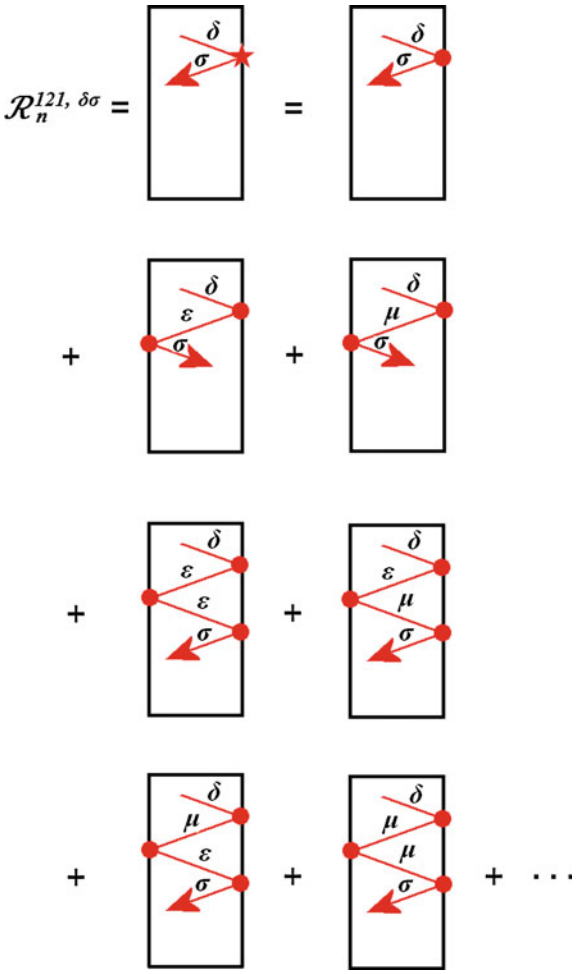


Fig. 8 Terms of the multiple scattering internal reflection amplitude $\mathcal{R}_n^{121,\delta\sigma}$ with the polarization state δ at the beginning and σ at the end

The other multiple-scattering internal reflection amplitudes $\mathcal{R}_n^{121,\mu\mu}$ and $\mathcal{R}_n^{121,\mu\varepsilon}$ are given by the same expressions with $\varepsilon \leftrightarrow \mu$.

Using the derivation of Sect. 4, the Debye series decomposition of the partial wave scattering amplitudes a_n, b_n, q_n is found to be

$$a_n = (1/2)\left[1 - R_n^{212,\varepsilon\varepsilon} + \sum_{\nu} T_n^{21,\varepsilon\nu} T_n^{12,\nu\varepsilon} + \sum_{\nu}\sum_{\tau} T_n^{21,\varepsilon\nu} \mathcal{R}_n^{121,\nu\tau} T_n^{12,\tau\varepsilon}\right] \tag{8.23a}$$

$$b_n = (1/2)\left[1 - R_n^{212,\mu\mu} + \sum_{\nu} T_n^{21,\mu\nu} T_n^{12,\nu\mu} + \sum_{\nu}\sum_{\tau} T_n^{21,\mu\nu} \mathcal{R}_n^{121,\nu\tau} T_n^{12,\tau\mu}\right] \tag{8.23b}$$

$$q_n = (-1/2)\left[-R_n^{212,\mu\varepsilon} + \sum_{\nu} T_n^{21,\mu\nu} T_n^{12,\nu\varepsilon} + \sum_{\nu}\sum_{\tau} T_n^{21,\mu\nu} \mathcal{R}_n^{121,\nu\tau} T_n^{12,\tau\varepsilon}\right] \tag{8.23c}$$

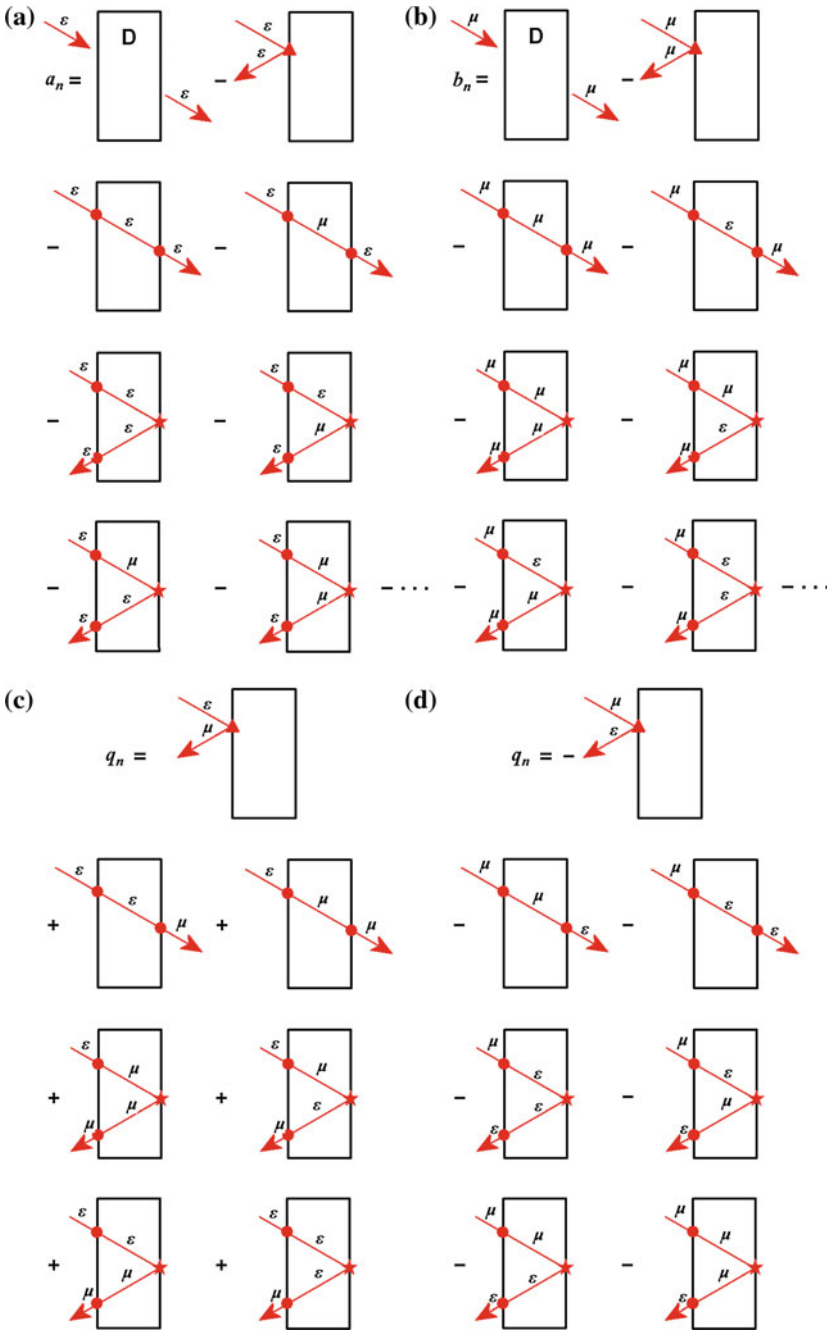
$$= (1/2)\left[-R_n^{212,\varepsilon\mu} + \sum_{\nu} T_n^{21,\varepsilon\nu} T_n^{12,\nu\mu} + \sum_{\nu}\sum_{\tau} T_n^{21,\varepsilon\nu} \mathcal{R}_n^{121,\nu\tau} T_n^{12,\tau\mu}\right]. \tag{8.23d}$$

The individual Debye series terms for a_n and b_n either keep the partial wave in the same polarization state from the time it enters the cylinder to the time it finally exits it, or the terms can switch the polarization state of the partial wave at any one of the interactions with the surface, as long as the polarization eventually switches back before the partial wave leaves the cylinder. The individual Debye series terms for the cross-polarized partial wave scattering amplitude q_n switch the polarization state of the partial wave at any of its interactions with the surface. The terms of Eqs. (8.23a–8.23d) are pictorially illustrated in Fig. 9.

As an application of the Debye series for the diagonal incidence geometry, we consider the evolution and extinguishing of the first-order rainbow as the tilt angle ζ of the incident plane wave progressively increases from normal incidence, with the external medium being air having $m_2 = 1$. In ray theory, the path of a ray through the cylinder can either be viewed from the side where one sees the projection of the path in the ρz plane, or from the top where one sees the projection of the path in the xy plane as in Fig. 2a, b of Adler et al. (1997). As seen from the top, changing the tilt angle of the incident plane wave is equivalent to changing the effective refractive index m' of the cylinder according to

$$m' = m_1 \cos(\chi)/\cos(\zeta). \tag{8.24}$$

Let φ_i be the angle of incidence of the projection of the incoming ray as seen in the horizontal xy plane, where φ_i is positive when the y coordinate of the incident ray is positive and φ_i is negative when y is negative [see Fig. 2a of Wait (1955)]. The transmitted angle φ_t is given by Snell's law


(a)
a_n =
D
(b)
b_n =
D
(c)
q_n =
(d)
q_n = –

◀ **Fig. 9** Terms of the polarization-preserving partial wave scattering amplitude a_n with polarization state ε at the beginning and ε at the end, b_n with the polarization state μ at the beginning and μ at the end, and the polarization-changing scattering amplitude q_n with the polarization state ε (or μ) at the beginning and μ (or ε) at the end. D denotes diffraction

$$\sin(\varphi_i) = m' \sin(\varphi_t), \tag{8.25}$$

and the positive or negative deflection angle of the ray after one internal reflection is

$$\theta = \pi + 2\varphi_i - 4\varphi_t. \tag{8.26}$$

The Descartes angle of the first-order rainbow θ^{R} is given by Eq. (8.26) with

$$\cos\left(\varphi_i^R\right) = \left[\left(m'^2 - 1\right)/3\right]^{1/2}. \tag{8.27}$$

The first-order rainbow for scattering of a plane wave by a *sphere* is a 2-ray-to-0-ray transition as a function of the deflection angle θ. For positive deflection angles, two supernumerary rays interfere for $\theta > \theta^{\mathrm{R}}$, and no rays occur in the evanescent complex-ray region $\theta < \theta^{\mathrm{R}}$. However, for scattering of a plane wave at normal incidence by a *cylinder*, two branches of the first-order rainbow occur to either side of back-scattering, one for incident rays with $\varphi_i^{\mathrm{R}} > 0$ and the other for $\varphi_i^{\mathrm{R}} < 0$. As is illustrated in Fig. 10, when $m' = 1.333$ at normal incidence, the Descartes rays with $\varphi_i^{\mathrm{R}} = \pm 59.41°$ give a pair of rainbows at $\theta^{\mathrm{R}} = \pm\ 137.92°$. Each of the rainbows is a 2-ray-to-0-ray transition. It is also useful to keep track of the deflection angle θ^{g} of the grazing incidence rays with $\varphi_i = \pm\ 90°$. When $m' = 1.333$ at normal incidence, $\varphi_i = \pm 90°$ gives $\theta^{\mathrm{g}} = \pm\ 165.64°$ for the grazing rays. Figure 10 shows that for this value of refractive index there are no $p = 2$ rays at any deflection angle in the interval $0° \le |\theta| < |\theta^{\mathrm{R}}| = 137.92°$, there are two supernumerary rays at each angle in the interval $|\theta^{\mathrm{R}}| < |\theta| < |\theta^{\mathrm{g}}| = 165.64°$ illustrating the 2-ray-to-0-ray transition, and there is one ray at each angle in the interval $|\theta^{\mathrm{g}}| < |\theta| \le 180°$.

Since varying the tilt angle ζ of the incident plane wave varies the effective refractive index m' of the cylinder for the projection of the rays in the xy plane, the deflection angle of the rainbow will evolve as a function of tilt. When the tilt of the incident plane wave is such that m' has increased to 1.414, the grazing incidence rays are back-scattered with $\theta^{\mathrm{g}} = \pm\ 180°$. So the angular interval $0° \le |\theta| < |\theta^{\mathrm{R}}| = 148.39°$ contains no rays, and the interval $\theta^{\mathrm{R}} < |\theta| \le 180°$ contains two supernumerary rays at each deflection angle. For a larger value of the tilt such that $m' > 1.414$, the deflection angle of the $\varphi_i = +90°$ and $-90°$ grazing incident rays becomes greater than 180° and less than −180°, respectively; this is also indicated in Fig. 10. As a result, the deflection angle interval $0° \le |\theta| < |\theta^{\mathrm{R}}|$

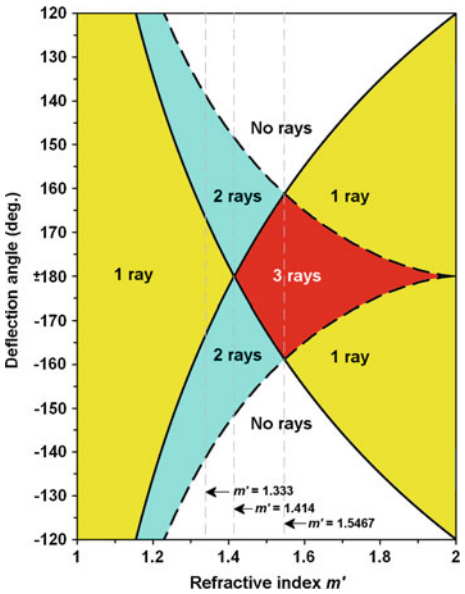


Fig. 10 Number of $p = 2$ scattered rays at each deflection angle in various regions for scattering of a tilted plane wave by a cylinder, as a function of the cylinder's effective refractive index m'. The *solid lines* indicate the scattering trajectory of the pair of grazing incident rays, and the *dashed lines* indicate the scattering trajectory of the two branches of the $p = 2$ rainbow

contains no rays, the interval $|\theta^R| < |\theta| < |\theta^g|$ contains two supernumerary rays at each angle with the rainbow again being a 2-ray-to-0-ray transition, and the interval $|\theta^g| < |\theta| \leq 180°$ now contains three rays at each deflection angle.

As the plane wave tilt and m' increase further, the 3-ray interval grows while the 2-ray interval $|\theta^R| < |\theta| < |\theta^g|$ shrinks, until at $m' \approx 1.5467$ the deflection angle of the positive and negative grazing incident rays coincides with that of the $\varphi_i^R < 0$ and $\varphi_i^R > 0$ Descartes rainbow ray at $\theta = \pm 161.12°$. The angular interval $0° \leq |\theta| < |\theta^R| = |\theta^g|$ now contains no rays, and the angular interval between the two rainbow branches contains three rays. As is also indicated in Fig. 10, when m' increases yet further, there are no rays in the angular interval $0° \leq |\theta| < |\theta^g|$, there is one ray at each deflection angle in the interval $|\theta^g| < |\theta| < |\theta^R|$, and there are three rays at each angle in the interval $|\theta^R| < |\theta| \leq 180°$. The pair of $p = 2$ rainbows for positive and negative φ_i^R have evolved into a pair of 3-ray-to-1-ray transitions associated with the cusp caustic, as is shown in Fig. 13 of Berry et al. (1979) and Fig. 2.8 of Berry and Upstill (1980). Finally, when $m' = 2$ both branches of the rainbow occur at $\theta^R = \pm 180°$. If m' is increased yet further, there ceases to be a one internal reflection rainbow; it has been extinguished. No rays occur in the angular interval $0° \leq |\theta| < |\theta^g|$, and there is one ray at each deflection angle in the interval $|\theta^g| < |\theta| \leq 180°$.

For a glass rod with refractive index $m_1 = 1.484$, the effective refractive index becomes $m' = 1.5467$ when the tilt angle is $\zeta = 21.68°$ and the rainbow extinction transition (e) at $m' = 2$ occurs when $\zeta_e = 50.72°$. This transition for a high-quality glass rod has been photographed in the laboratory, and the results are shown in Fig. 8 of Lock and Adler (1997). The Debye series amplitudes

$$a_n = (-1/2)\left(T_n^{21,\varepsilon\varepsilon} R_n^{121,\varepsilon\varepsilon} T_n^{12,\varepsilon\varepsilon} + T_n^{21,\varepsilon\varepsilon} R_n^{121,\varepsilon\mu} T_n^{12,\mu\varepsilon} + T_n^{21,\varepsilon\mu} R_n^{121,\mu\varepsilon} T_n^{12,\varepsilon\varepsilon} + T_n^{21,\varepsilon\mu} R_n^{121,\mu\mu} T_n^{12,\mu\varepsilon}\right) \tag{8.28a}$$

$$b_n = (-1/2)\left(T_n^{21,\mu\mu} R_n^{121,\mu\mu} T_n^{12,\mu\mu} + T_n^{21,\mu\mu} R_n^{121,\mu\varepsilon} T_n^{12,\varepsilon\mu} + T_n^{21,\mu\varepsilon} R_n^{121,\varepsilon\mu} T_n^{12,\mu\mu} + T_n^{21,\mu\varepsilon} R_n^{121,\varepsilon\varepsilon} T_n^{12,\varepsilon\mu}\right) \tag{8.28b}$$

$$q_n = (-1/2)\left(T_n^{21,\varepsilon\mu} R_n^{121,\mu\mu} T_n^{12,\mu\mu} + T_n^{21,\varepsilon\mu} R_n^{121,\mu\varepsilon} T_n^{12,\varepsilon\mu} + T_n^{21,\varepsilon\varepsilon} R_n^{121,\varepsilon\mu} T_n^{12,\mu\mu} + T_n^{21,\varepsilon\varepsilon} R_n^{121,\varepsilon\varepsilon} T_n^{12,\varepsilon\mu}\right) \tag{8.28c}$$

were also computed for $m_1 = 1.484$, $m_2 = 1$, $x = 2\pi a/\lambda = 1000$, and $\zeta = 39°$, 45°, 50.72°, and 56°. The resulting scattered intensity is shown in Fig. 4 of Lock and Adler (1997). The rainbow's 3-ray-to-1-ray transition for $\zeta = 39°$, 45° and the extinguishing of the rainbow at $\zeta \approx 50.72°$ are clearly evident.

9 Debye Series for a Scattering by a Spheroid

In this section, we consider scattering of a plane wave by a homogeneous prolate spheroid of refractive index m_1, semi-major axis a, semi-minor axis b with $b < a$, which is in an exterior medium of refractive index m_2. The z axis coincides with the major axis of the spheroid, and the propagation direction of the incident plane wave lies in the xz plane making an angle ζ with respect to the positive z axis. The case of $\zeta = 0°$ where the plane wave is incident along the -z axis is known as end-on incidence, and $\zeta = 90°$ where the plane wave is incident along the -x axis is known as side-on incidence (Asano and Yamamoto 1975). As in Sect. 8, the plane wave is ε-polarized if its electric field lies in the xy plane, and it is μ-polarized if its magnetic field lies in the xy plane. Because of the spheroid geometry's inherent breaking of spherical symmetry, we will find that every interaction of an electromagnetic wave with the spheroid surface again contains both polarization-preserving and polarization-changing components (Lock 1996). In addition, each partial wave approaching the surface for both the scalar wave and electromagnetic wave cases will be coupled to all the partial waves leaving the surface. We first address the scattering problem for scalar waves in order to focus on the partial wave coupling. We then briefly summarize the electromagnetic problem to include the polarization-changing behavior.

9.1 Scattering of Scalar Waves

This scattering problem is traditionally solved using spheroidal wave functions (Flammer 1957). A point in space (x, y, z) is expressed in terms of prolate spheroidal coordinates by

$$x = f\sinh(u)\sin(v)\cos(\varphi) \tag{9.1a}$$

$$y = f\sinh(u)\sin(v)\sin(\varphi) \tag{9.1b}$$

$$z = f\cosh(u)\cos(v) \tag{9.1c}$$

where

$$f = \left(a^2 - b^2\right)^{1/2} \tag{9.2}$$

and $0 \le u < \infty$, $0 \le v \le \pi$, and $0 \le \varphi \le 2\pi$. The usual convention for prolate spheroidal coordinates is to use (ξ, η, φ) where $\xi = \cosh(u)$ is the radial-like coordinate with $1 \le \xi < \infty$, and $\eta = \cos(v)$ is the polar-like coordinate with $-1 \le \eta \le 1$ (Flammer 1957). Surfaces of constant ξ are confocal spheroids, and surfaces of constant η are confocal hyperboloids. The particle surface is $\xi_0 = a/f$ for all η and φ.

The solution to the scalar wave equation in terms of prolate spheroidal wave functions is

$$\psi^{(K)}\left(m_j kf; \xi, \eta, \varphi\right) = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} F_{m,n} R_{m,n}^{(K)}\left(m_j kf; \xi\right) S_{m,n}\left(m_j kf; \eta\right)\exp(\mathrm{i}m\varphi) \tag{9.3}$$

where $F_{m,n}$ are complex constants. The spheroidal angular functions $S_{m,n}(m_j kf; \eta)$ depend on both the spheroidal angular variable η and $m_j kf$, and reduce in the $f \to 0$ limit to associated Legendre functions $P_n^m[\cos(\theta)]$. The angular functions $S_{m,n}(m_j kf; \eta)$ and $S_{m,n'}(m_j kf; \eta)$ for the same refractive index m_j are orthogonal to each other when integrated over the complete range of η. The spheroidal radial functions of the K-type, $R_{m,n}^{(K)}\left(m_j kf; \xi\right)$ for $K = 1,2,3,4$, depend on both the spheroidal radial variable ξ and $m_j kf$, and reduce for $\xi = r/f$ in the limit $f \to 0$ to $j_n(m_j kr)$, $n_n(m_j kr)$ $h_n^{(1)}\left(m_j kr\right)$, $h_n^{(2)}\left(m_j kr\right)$, respectively. The radial functions satisfy the Wronskian relation

$$R_{m,n}^{(1)}\left(m_j kf; \xi\right)R_{m,n}^{(2)\prime}\left(m_j kf; \xi\right) - R_{m,n}^{(1)\prime}\left(m_j kf; \xi\right)R_{m,n}^{(2)}\left(m_j kf; \xi\right) = 1/\left[m_j kf\left(\xi^2 - 1\right)\right], \tag{9.4}$$

independent of m and n, where R' is the derivative of R with respect to ξ. As a side note, the convention of writing the spheroidal wave-functions with the subscripts in the order m, n rather than n, m dates back to at least the time of Stratton [see pp. 420–421 of Stratton et al. (1941)].

The scalar wave scattering problem is posed and solved in the usual way. The fields of the incident, scattered, and interior waves are written in terms of spheroidal wave functions with the incident beam, scattered wave, and interior wave coefficients $m_2\, Z_{m,n,} - m_2\, b_{m,n}$ and $m_1\, d_{m,n}$, respectively. The angle of incidence of the plane wave is encoded in the beam shape coefficients $Z_{m,n}$ [see p. 48 of Flammer (1957)]. The incident and scattered wave functions depend on m_2kf and ξ,η,φ, while the interior wave function depends on m_1kf and ξ,η,φ. The boundary conditions for the continuity of the wave function and its derivative with respect to ξ are applied at the surface of the spheroid. The properties of the azimuthal functions $\exp(im\varphi)$ insure that the m-dependence is identical for the incident, scattered, and interior waves. But since the angular functions depend on m_2kf and m_1kf, the functions $S_{m,n}(m_2kf;\ \eta)$ and $S_{m,n'}\ (m_1kf;\ \eta)$ for different refractive indices m_1 and m_2 are not orthogonal to each other when integrated over the complete range of η. This lack of orthogonality causes each partial wave n' of the incident beam to be coupled to all the partial waves n of the scattered and interior waves. The strength of the coupling is given by the matrix elements

$$C_{m,N1,N2}(M_1kf, M_2kf) = \int_{-1}^{1} \mathrm{d}\eta\, S_{m,N1}(M_1kf;\eta) S_{m,N2}(M_2kf;\eta) \Big/ \int_{-1}^{1} \mathrm{d}\eta \left[S_{m,N1}(M_1kf;\eta)\right]^2, \tag{9.5}$$

where N_1 and N_2 are partial waves and M_1 and M_2 are refractive indices. If the largest partial wave considered is L, boundary condition matching produces $2L$ coupled algebraic equations in the $2L$ partial wave scattering and interior amplitudes that can be solved numerically as long as L is not too large.

The quantities being considered can be expressed in terms of eight fundamental amplitude matrices in the partial wave number indices N1 and N2 in analogy to Eqs. (2.15a–2.15d) for a homogeneous sphere

$$\begin{aligned} N_{m,N1,N2}(M_1kf, M_2kf;\xi_0) = C_{m,N1,N2}(M_1kf, M_2kf) \Big[& R^{(1)}_{m,N1}(M_1kf;\xi_0) R^{(1)\prime}_{m,N2}(M_2kf;\xi_0) \\ & - R^{(1)\prime}_{m,N1}(M_1kf;\xi_0) R^{(1)}_{m,N2}(M_2kf;\xi_0) \Big] \end{aligned} \tag{9.6a}$$

$$\begin{aligned} D_{m,N1,N2}(M_1kf, M_2kf;\xi_0) = C_{m,N1,N2}(M_1kf, M_2kf) \Big[& R^{(2)}_{m,N1}(M_1kf;\xi_0) R^{(1)\prime}_{m,N2}(M_2kf;\xi_0) \\ & - R^{(2)\prime}_{m,N1}(M_1kf;\xi_0) R^{(1)}_{m,N2}(M_2kf;\xi_0) \Big] \end{aligned} \tag{9.6b}$$

$$P_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = C_{m,N1,N2}(M_1kf, M_2kf)\Big[R^{(1)}_{m,N1}(M_1kf; \xi_0)R^{(2)\prime}_{m,N2}(M_2kf; \xi_0) - R^{(1)\prime}_{m,N1}(M_1kf; \xi_0)R^{(2)}_{m,N2}(M_2kf; \xi_0)\Big] \tag{9.6c}$$

$$Q_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = C_{m,N1,N2}(M_1kf, M_2kf)\Big[R^{(2)}_{m,N1}(M_1kf; \xi_0)R^{(2)\prime}_{m,N2}(M_2kf; \xi_0) - R^{(2)\prime}_{m,N1}(M_1kf; \xi_0)R^{(2)}_{m,N2}(M_2kf; \xi_0)\Big] \tag{9.6d}$$

and

$$\underline{N}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = C_{m,N1,N2}(M_1kf, M_2kf)\Big[R^{(1)}_{m,N2}(M_2kf; \xi_0)R^{(1)\prime}_{m,N1}(M_1kf; \xi_0) - R^{(1)\prime}_{m,N2}(M_2kf; \xi_0)R^{(1)}_{m,N1}(M_1kf; \xi_0)\Big] \tag{9.7a}$$

$$\underline{D}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = C_{m,N1,N2}(M_1kf, M_2kf)\Big[R^{(2)}_{m,N2}(M_2kf; \xi_0)R^{(1)\prime}_{m,N1}(M_1kf; \xi_0) - R^{(2)\prime}_{m,N2}(M_2kf; \xi_0)R^{(1)}_{m,N1}(M_1kf; \xi_0)\Big] \tag{9.7b}$$

$$\underline{P}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = C_{m,N1,N2}(M_1kf, M_2kf)\Big[R^{(1)}_{m,N2}(M_2kf; \xi_0)R^{(2)\prime}_{m,N1}(M_1kf; \xi_0) - R^{(1)\prime}_{m,N2}(M_2kf; \xi_0)R^{(2)}_{m,N1}(M_1kf; \xi_0)\Big] \tag{9.7c}$$

$$\underline{Q}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = C_{m,N1,N2}(M_1kf, M_2kf)\Big[R^{(2)}_{m,N2}(M_2kf; \xi_0)R^{(2)\prime}_{m,N1}(M_1kf; \xi_0) - R^{(2)\prime}_{m,N2}(M_2kf; \xi_0)R^{(2)}_{m,N1}(M_1kf; \xi_0)\Big]. \tag{9.7d}$$

It should be noted that the difference between the underlined matrices of Eqs. (9.7a–9.7d) and the non-underlined matrices of Eqs. (9.6a–9.6d) is the interchange of the partial wave subscripts N1 ↔ N2 and refractive indices M1 ↔ M2 in the radial functions. Further, it will prove useful to also make use of the following combinations of the partial wave amplitude matrices in analogy to Eqs. (2.19, 3.5a, 3.5b, 3.8) for a homogeneous sphere,

$$U_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = N_{m,N1,N2}(M_1kf, M_2kf; \xi_0) + \mathrm{i}\, D_{m,N1,N2}(M_1kf, M_2kf; \xi_0) \tag{9.8a}$$

$$M_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = \big[N_{m,N1,N2}(M_1kf, M_2kf; \xi_0) + Q_{m,N1,N2}(M_1kf, M_2kf; \xi_0)\big] + \mathrm{i}\big[D_{m,N1,N2}(M_1kf, M_2kf; \xi_0) - P_{m,N1,N2}(M_1kf, M_2kf; \xi_0)\big] \tag{9.8b}$$

$$L_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = -\left[N_{m,N1,N2}(M_1kf, M_2kf; \xi_0) - Q_{m,N1,N2}(M_1kf, M_2kf; \xi_0)\right] - \mathrm{i}\left[D_{m,N1,N2}(M_1kf, M_2kf; \xi_0) + P_{m,N1,N2}(M_1kf, M_2kf; \xi_0)\right] \tag{9.8c}$$

and

$$\underline{U}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = \underline{N}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) + \mathrm{i}\underline{D}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) \tag{9.9a}$$

$$\underline{M}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = \left[\underline{N}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) + \underline{Q}_{m,N1,N2}(M_1kf, M_2kf; \xi_0)\right] + i\left[\underline{D}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) - \underline{P}_{m,N1,N2}(M_1kf, M_2kf; \xi_0)\right] \tag{9.9b}$$

$$\underline{J}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) = -\left[\underline{N}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) - \underline{Q}_{m,N1,N2}(M_1kf, M_2kf; \xi_0)\right] + \mathrm{i}\left[\underline{D}_{m,N1,N2}(M_1kf, M_2kf; \xi_0) + \underline{P}_{m,N1,N2}(M_1kf, M_2kf; \xi_0)\right]. \tag{9.9c}$$

In analogy to Eqs. (2.18, 2.21) for a homogeneous sphere, the partial wave scattering and interior amplitudes are found to be

$$b_{m,n} = \sum_{n''}\sum_{n'}\left[\underline{U}^{-1}(m_1kf, m_2kf; \xi_0)\right]_{m,n,n''}\underline{N}_{m,n'',n'}(m_1kf, m_2kf; \xi_0)Z_{m,n'} \tag{9.10a}$$

$$d_{m,n} = -\left[\mathrm{i}/m_1kf\left(\xi_0^2 - 1\right)\right]\sum_{n'}\left[U^{-1}(m_2kf, m_1kf; \xi_0)\right]_{m,n,n'}Z_{m,n'}. \tag{9.10b}$$

where U^{-1} and $\underline{U}^{-1}$ are the inverse of the matrices U and $\underline{U}$. The order of the matrices reads from right to left. In Eq. (9.10a) the partial wave number of the incident beam shape coefficient Z is n', then the $\underline{N}$ matrix couples n' to the intermediate partial wave n'', and finally the $\underline{U}^{-1}$ matrix couples n'' to the scattered partial wave n.

Similarly, the partial wave transmission and reflection amplitude matrices can be determined for the output partial wave n when a single partial wave n' is incident on the spheroid surface from either the external medium or the particle interior. In analogy to Eqs. (3.4a, 3.7a, 3.4b, 3.7b) for a homogeneous sphere one obtains

$$T^{21}_{m,n,n'} = -\left[2\mathrm{i}/m_2kf\left(\xi_0^2 - 1\right)\right]\left[M^{-1}(m_2kf, m_1kf; \xi_0)\right]_{m,n,n'} \tag{9.11a}$$

$$\underline{T}^{12}_{m,n,n'} = -\left[2\mathrm{i}/m_1kf\left(\xi_0^2 - 1\right)\right]\left[\underline{M}^{-1}(m_1kf, m_2kf; \xi_0)\right]_{m,n,n'} \tag{9.11b}$$

$$\underline{R}^{212}_{m,n,n'} = \sum_{n''} \left[\underline{M}^{-1}(m_1 kf, m_2 kf; \xi_0)\right]_{m,n,n''} \underline{J}_{m,n'',n'}(m_1 kf, m_2 kf; \xi_0) \tag{9.11c}$$

$$R^{121}_{m,n,n'} = \sum_{n''} \left[M^{-1}(m_2 kf, m_1 kf; \xi_0)\right]_{m,n,n''} L_{m,n'',n'}(m_2 kf, m_1 kf; \xi_0) \tag{9.11d}$$

The superscripts in Eqs. (9.11a–9.11d) describing medium 1 or 2 read from left to right as usual. But since the transmission and reflection amplitudes are now matrices rather than scalars, the partial wave number subscripts read from right to left. In addition, the difference in the physical meaning of the underlined and non-underlined matrices now becomes apparent. The partial wave in question ends up inside the spheroid for non-underlined amplitude matrices, and it ends up outside the spheroid for the underlined matrices. Finally, extending the derivation outlined in Sect. 4 to matrices, we find in analogy to Eqs. (4.17, 4.13) for a homogeneous sphere that Debye series of the scattering and interior amplitudes for scalar waves is

$$b_{m,n} = (1/2) \sum_{n'} \left\{ \delta_{n,n'} - \underline{R}^{212}_{m,n,n'} - \sum_{n''} \sum_{n'''} \underline{T}^{12}_{m,n,n''} \left[(I - R^{121})^{-1} \right]_{m,n'',n'''} T^{21}_{m,n''',n'} \right\} Z_{m,n'} \tag{9.12a}$$

$$d_{m,n} = \sum_{n''} \sum_{n'} \left[(I - R^{121})^{-1} \right]_{m,n,n''} T^{21}_{m,n'',n'} Z_{m,n'}, \tag{9.12b}$$

where I is the identity matrix whose elements are given by the Kronecker delta, $\delta_{n,n'}$. It should again be noted for emphasis that since Eqs. (9.12a, 9.12b) involve matrix multiplications, the ordering of the R and T factors reads from right to left, rather than left to right as was the case in all the previous sections. Also, the progression of the partial wave number in the last term of Eq. (9.12a) is $n' \to n''' \to n'' \to n$. As an example, the partial wave coupling of all the incident partial waves n' to a single partial wave n via the intermediate partial wave n'' for the $\underline{T}^{12}\, T^{21}$ term is pictorially illustrated in Fig. 11.

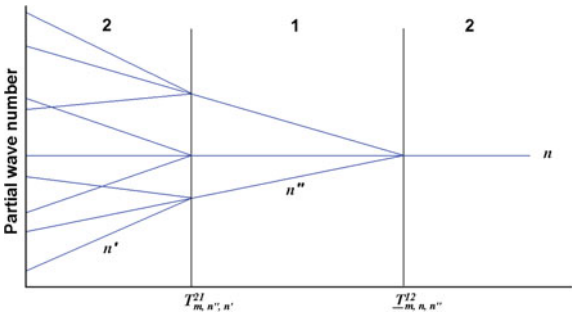


Fig. 11 For transmission through a spheroid all incident partial waves n', reading from left to right in the figure, are coupled to each partial wave n'' inside, and all partial waves n'' are coupled to each partial wave n that is transmitted back out

9.2 Scattering of Electromagnetic Waves

Turning now to scattering of electromagnetic waves, the solution to the vector wave equation is written in terms of spheroidal coordinates ξ, η, φ, vector spheroidal functions, and spheroidal coordinate unit vectors $\mathbf{u}_\xi$, $\mathbf{u}_\eta$, and $\mathbf{u}_\varphi$ as [see p. 70 of Flammer (1957)]

$$\mathbf{M}_{m,n}^{(K)}\left(m_j kf;\xi,\eta,\varphi\right) = \nabla\psi_{m,n}^{(K)}\left(m_j kf;\xi,\eta,\varphi\right)\times\mathbf{r} \tag{9.13a}$$

$$\mathbf{N}_{m,n}^{(K)}\left(m_j kf;\xi,\eta,\varphi\right) = \left(1/m_j k\right)\nabla\times\mathbf{M}_{m,n}^{(K)}\left(m_j kf;\xi,\eta,\varphi\right). \tag{9.13b}$$

An ε-polarized or μ-polarized incident plane wave can be written as a sum over these functions with the beam shape coefficients $m_2\, G_{m,n'}^{\varepsilon}$ and $m_2\, G_{m,n'}^{\mu}$ [see p. 71 of Flammer (1957) and Eqs. (40, 41) of Asano and Yamamoto (1975)]. Similarly, the scattered wave has the partial wave scattering amplitudes $-m_2\, a_{m,n}$ for the μ polarization and $-m_2\, b_{m,n}$ for the ε polarization, which possess polarization-preserving and polarization changing components given by

$$a_{m,n} = \mathrm{i}\sum_{n'} A_{m,n,n'}^{\mu\varepsilon} G_{m,n'}^{\varepsilon} + \sum_{n'} A_{m,n,n'}^{\mu\mu} G_{m,n'}^{\mu} \tag{9.14a}$$

$$b_{m,n} = \mathrm{i}\sum_{n'} B_{m,n,n'}^{\varepsilon\varepsilon} G_{m,n'}^{\varepsilon} + \sum_{n'} B_{m,n,n'}^{\varepsilon\mu} G_{m,n'}^{\mu}. \tag{9.14b}$$

Since these quantities are matrices, their order reads from right to left, as do the partial wave subscripts $n' \to n$ and the polarization state superscripts. The same is true of the interior field coefficients $m_1\, c_{m,n}$ and $m_1\, d_{m,n,}$

$$c_{m,n} = \mathrm{i}\sum_{n'} C_{m,n,n'}^{\mu\varepsilon} G_{m,n'}^{\varepsilon} + \sum_{n'} C_{m,n,n'}^{\mu\mu} G_{m,n'}^{\mu} \tag{9.15a}$$

$$d_{m,n} = \mathrm{i}\sum_{n'} D_{m,n,n'}^{\varepsilon\varepsilon} G_{m,n'}^{\varepsilon} + \sum_{n'} D_{m,n,n'}^{\varepsilon\mu} G_{m,n'}^{\mu}. \tag{9.15b}$$

As was the case in Sect. 8, if a single ε-polarized or μ-polarized partial wave were incident on the spheroid, it is coupled to the transmission and reflection amplitudes of all other partial waves with both polarization-preserving and polarization-changing components. The detailed expressions for these amplitudes are given in Xu et al. (2010a). Using the matrix version of the procedure of Sect. 4, the Debye series for the interior fields turns out to be

$$C^{\mu\varepsilon}_{m,n,n'} = T^{21,\mu\varepsilon}_{m,n,n'} + \sum_{\delta}\sum_{n''} R^{121,\mu\delta}_{m,n,n''} T^{21,\delta\varepsilon}_{m,n'',n'} + \sum_{\delta}\sum_{\sigma}\sum_{n''}\sum_{n'''} R^{121,\mu\delta}_{m,n,n''} R^{121,\delta\sigma}_{m,n'',n'''} T^{21,\sigma\varepsilon}_{m,n''',n'} + \cdots \tag{9.16a}$$

$$C^{\mu\mu}_{m,n,n'} = T^{21,\mu\mu}_{m,n,n'} + \sum_{\delta}\sum_{n''} R^{121,\mu\delta}_{m,n,n''} T^{21,\delta\mu}_{m,n'',n'} + \sum_{\delta}\sum_{\sigma}\sum_{n''}\sum_{n'''} R^{121,\mu\delta}_{m,n,n''} R^{121,\delta\sigma}_{m,n'',n'''} T^{21,\sigma\mu}_{m,n''',n'} + \cdots \tag{9.16b}$$

$$D^{\varepsilon\varepsilon}_{m,n,n'} = T^{21,\varepsilon\varepsilon}_{m,n,n'} + \sum_{\delta}\sum_{n'} R^{12,\varepsilon\delta}_{m,n,n''} T^{21,\delta\varepsilon}_{m,n'',n'} + \sum_{\delta}\sum_{\sigma}\sum_{n''}\sum_{n'''} R^{121,\varepsilon\delta}_{m,n,n''} R^{121,\delta\sigma}_{m,n'',n'''} T^{21,\sigma\varepsilon}_{m,n''',n'} + \cdots \tag{9.16c}$$

$$D^{\varepsilon\mu}_{m,n,n'} = T^{21,\varepsilon\mu}_{m,n,n'} + \sum_{\delta}\sum_{n''} R^{121,\varepsilon\delta}_{m,n,n''} T^{21,\delta\mu}_{m,n'',n'} + \sum_{\delta}\sum_{\sigma}\sum_{n''}\sum_{n'''} R^{121,\varepsilon\delta}_{m,n,n''} R^{121,\delta\sigma}_{m,n'',n'''} T^{21,\sigma\mu}_{m,n''',n'} + \cdots, \tag{9.16d}$$

where all partial waves and all polarization states are summed over at every interaction with the surface. Using these results, the Debye series for the scattering amplitudes is then found to be

$$A^{\mu\mu}_{m,n,n'} = (1/2)\left[\delta_{n,n'} - \underline{R}^{212,\mu\mu}_{m,n,n'} - \sum_{n''}\left(\underline{T}^{12,\mu\varepsilon}_{m,n,n''} D^{\varepsilon\mu}_{m,n'',n'} + \underline{T}^{12,\mu\mu}_{m,n,n''} C^{\mu\mu}_{m,n'',n'}\right)\right] \tag{9.17a}$$

$$A^{\mu\varepsilon}_{m,n,n'} = (1/2)\left[-\underline{R}^{212,\mu\varepsilon}_{m,n,n'} - \sum_{n''}\left(\underline{T}^{12,\mu\varepsilon}_{m,n,n''} D^{\varepsilon\varepsilon}_{m,n'',n'} + \underline{T}^{12,\mu\mu}_{m,n,n''} C^{\mu\varepsilon}_{m,n'',n'}\right)\right] \tag{9.17b}$$

$$B^{\varepsilon\varepsilon}_{m,n,n'} = (1/2)\left[\delta_{n,n'} - \underline{R}^{212,\varepsilon\varepsilon}_{m,n,n'} - \sum_{n''}\left(\underline{T}^{12,\varepsilon\varepsilon}_{m,n,n''} D^{\varepsilon\varepsilon}_{m,n'',n'} + \underline{T}^{12,\varepsilon\mu}_{m,n,n''} C^{\mu\varepsilon}_{m,n'',n'}\right)\right] \tag{9.17c}$$

$$B^{\varepsilon\mu}_{m,n,n'} = (1/2)\left[-\underline{R}^{212,\varepsilon\mu}_{m,n,n'} - \sum_{n''}\left(\underline{T}^{12,\varepsilon\varepsilon}_{m,n,n''} D^{\varepsilon\mu}_{m,n'',n'} + \underline{T}^{12,\varepsilon\mu}_{m,n,n''} C^{\mu\mu}_{m,n'',n'}\right)\right]. \tag{9.17d}$$

These expressions exhibit the same patterns that were found for scattering by a homogeneous sphere in Eqs. (4.18a), for a coated sphere in Eq. (7.11), for a multi-layer sphere in Eq. (7.28), for diagonal incidence on a cylinder in Eqs. (8.23a–8.23d), and for scalar wave scattering by a spheroid in Eq. (9.12a), now including both polarization coupling and partial wave coupling.

10 General Description of Time-Domain Scattering

The previous sections have implicitly assumed "steady-state" scattering of light with an intensity that is constant in time. But much more can be learned by considering scattering of a temporally short light pulse. For a given scattering angle θ, small-p Debye series terms such as diffraction-plus-reflection or transmission have a relatively short optical path length between the sphere's entrance and exit planes, while large-p terms such as transmission after three or four internal reflections have a relatively long optical path length between the entrance and exit planes. When a temporally short pulse is incident on a spherical particle, the small-p scattered pulses exit the sphere before the large-p scattered pulses do. In this way, scattering of a short pulse effects a separation of the various Debye series terms in the Lorenz–Mie scattered field (Méès et al. 2001a, b, 2009; Han et al. 2004; Bech and Leder 2004, 2006; Bakić et al. 2008; Laven 2010, 2011) without having to explicitly calculate them as was done in Sects. 4–9.

Consider an x-polarized incident plane wave pulse traveling in the $+z$ direction whose electric field is of the form $E^{\text{pulse}}(Z)\exp(\mathrm{i}k_0 Z)$ with the dominant wavelength λ_0, dominant wave number $k_0 = 2\pi/\lambda_0$, and $Z \equiv z - ct$. This pulse has the frequency-domain spectrum

$$A(k,k_0) = \int_{-\infty}^{\infty} \mathrm{d}Z\, E^{\text{pulse}}(Z)\exp(\mathrm{i}k_0 Z)\exp(-\mathrm{i}kZ). \tag{10.1}$$

If the Lorenz–Mie scattering response of the sphere to an x-polarized incident plane wave of wave number k and unit amplitude that is traveling in the $+z$ direction is $\mathbf{E}^{\mathbf{LM}}(k,\theta)$ as in Eq. (2.10a), then the composite frequency-domain spectrum of the scattered signal is $A(k, k_0)\ \mathbf{E}^{\mathbf{LM}}(k,\theta)$. The time-domain scattered field is then

$$\mathbf{E}^{\mathbf{scatt}}(\theta,t) = \int_{-\infty}^{\infty} \mathrm{d}k/(2\pi)A(k,k_0)\mathbf{E}^{\mathbf{LM}}(k,\theta)\exp(-\mathrm{i}ckt). \tag{10.2}$$

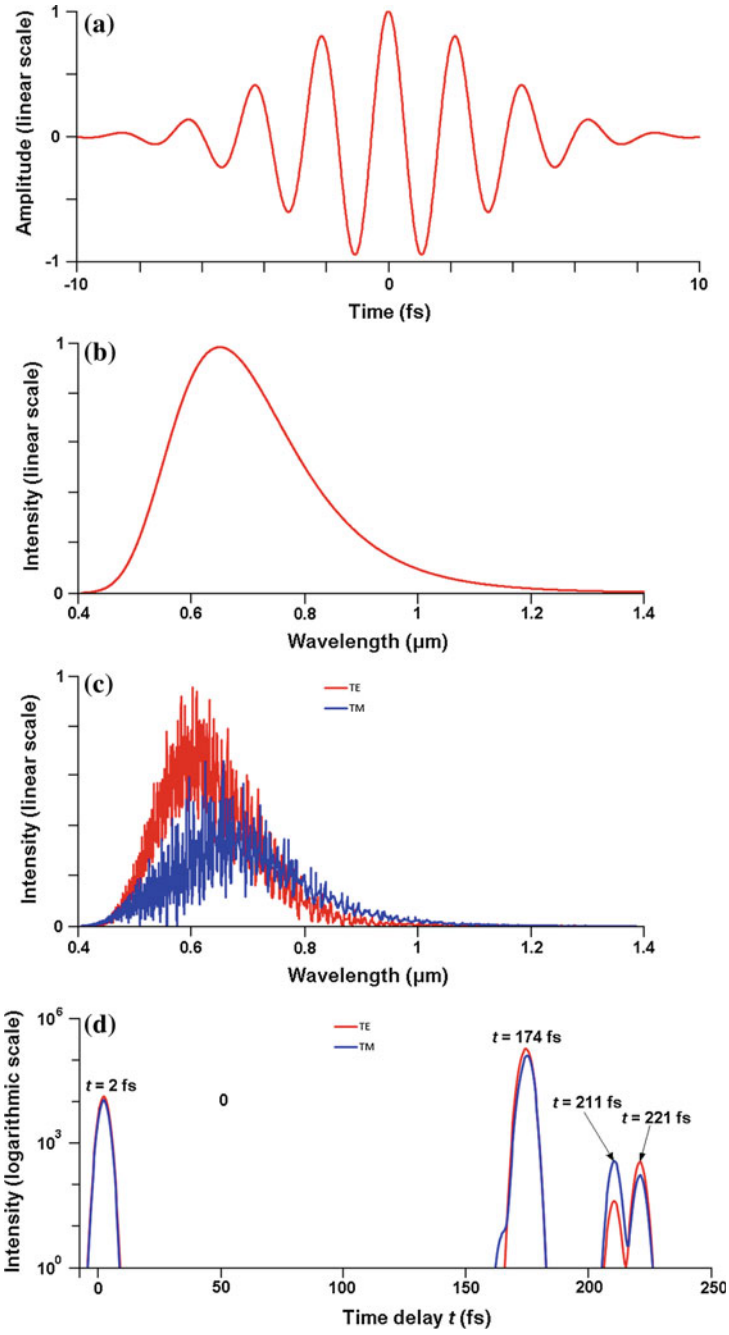
For the time-domain computations reported here and in Sects. 11, 12, we used the Gaussian incident plane wave pulse

$$\mathbf{E}^{\mathbf{pulse}}(Z) = E_0 \exp\left(-Z^2/\sigma^2\right)\mathbf{u}_{\mathbf{x}} \tag{10.3}$$

for which

$$A(k,k_0) = E_0\sigma(\pi)^{1/2}\exp\left[-(k-k_0)^2\sigma^2/4\right]. \tag{10.4}$$

Figure 12a shows the waveform of a Gaussian pulse of dominant wavelength $\lambda_0 = 0.65$ μm whose amplitude full-width at half-maximum is 5 fs. In this case


(a)
Amplitude (linear scale)
Time (fs)
(b)
Intensity (linear scale)
Wavelength (μm)
(c)
TE
TM
Intensity (linear scale)
Wavelength (μm)
(d)
TE
TM
t = 2 fs
t = 174 fs
t = 211 fs
t = 221 fs
Intensity (logarithmic scale)
Time delay t (fs)

◀ **Fig. 12** Scattering of a Gaussian electromagnetic pulse of nominal wavelength $\lambda_0 = 0.65$ μm and temporal duration 5 fs. The pulse waveform is shown in **a** and the magnitude-squared of its frequency-domain spectrum is shown in **b**. Lorenz–Mie calculations for scattering at $\theta = 150°$ of an x-polarized plane wave by a sphere with $a = 10$ μm, $m_1 = 1.33257$, and $m_2 = 1$ multiply the spectrum, producing the composite frequency-domain spectrum whose magnitude-squared is shown in **c**. Performing an inverse fast Fourier transform of the composite spectrum with a grid spacing of $\Delta t = 0.54$ fs generates the time-domain results shown in **d**

$\sigma = 0.90$ μm in Eq. (10.3), and the pulse consists of only a few cycles at the nominal wavelength. Using a $2^{15} = 32{,}768$ point Fast Fourier Transform, the time-domain waveform in Fig. 12a was converted into the frequency-domain spectrum, the square of the magnitude of this spectrum is shown in Fig. 12b as a function of wavelength. It has the bandwidth 0.573 μm $< \lambda <$ 0.751 μm measured at the −3 dB points, and 0.405 μm $< \lambda <$ 1.64 μm measured at the −60 dB points. As a preliminary example demonstrating the separation of the Lorenz–Mie scattered field in the time-domain, consider scattering at the angle $\theta = 150°$. The frequency-domain spectrum was truncated at the −60 dB points and Lorenz–Mie calculations were made at the remaining ∼10,000 evenly spaced wave numbers for which $A(k, k_0) \neq 0$. The Lorenz–Mie response at $\theta = 150°$ for each wave number was then multiplied by the frequency-domain spectrum function of Fig. 12b The magnitude-squared of the resulting composite spectrum is shown in Fig. 12c. Note that the composite spectrum of the scattered signals is significantly different for the TE and TM polarizations and, furthermore, neither polarization has its maximum exactly at $\lambda = 0.65$ μm, as might have been expected from the spectrum of the incident light as shown in Fig. 12b. More importantly, the very complicated ripples on the TE and TM composite spectra demonstrate the complexity of the scattering processes. Applying an inverse fast Fourier transform to the composite spectrum

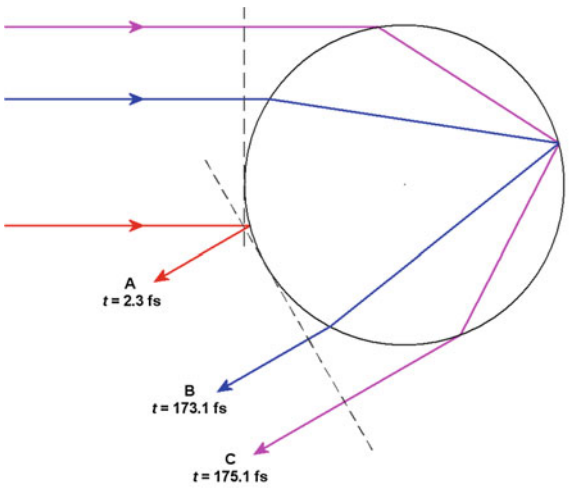


Fig. 13 Geometric rays contributing to scattering at $\theta = 150°$ for a sphere with $a = 10$ μm, $m_1 = 1.33257$, and $m_2 = 1$. The specified values of t refer to the propagation times between the entrance and exit planes indicated by the *dashed lines*

and taking the square of the magnitude of the result, the scattered intensity as a function of time at $\theta = 150°$ is obtained, and is shown in Fig. 12d.

Some understanding of Fig. 12d can be obtained by considering the geometric rays shown in Fig. 13 which result in scattering at $\theta = 150°$. The pulse in Fig. 12d at $t \approx 2$ fs corresponds to the externally reflected $p = 0$ ray A in Fig. 13 with a calculated time delay of $t = 2.3$ fs. The slightly misshapen pulse at $t \approx 174$ fs in Fig. 12d is due to the partial overlap of the two geometrical $p = 2$ rays B and C in Fig. 13 which have time delays of $t = 173.1$ fs and $t = 175.1$ fs. The remaining pulses in Fig. 12d are not due to geometrical rays. The pulse at $t \approx 211$ fs, for which the TM polarization is dominant, is caused by $p = 2$ surface waves, whereas the pulse at $t \approx 221$ fs is caused by the $p = 3$ complex ray associated with the second-order rainbow.

11 Time-Domain Scattering and the Ray Limit of the Debye Series

The time-domain results in Fig. 12 are useful, but they are limited to a single value of scattering angle θ. Much more information can be revealed by examining how the time-domain results vary as a function of θ. For example, the false-color map in Fig. 14 plots the scattered intensity as a function of t and θ (Laven 2010, 2011). Although this diagram has been calculated using Lorenz–Mie theory, the scattering contributions due to the individual terms of the Debye series are easily identifiable, as was mentioned at the beginning of Sect. 10. The time-domain Fourier transform of Eq. (10.2) essentially reads the interference structure of two or more p-processes in the scattered field occurring at the same scattering angle and separates them, without having to explicitly calculate the various Debye terms. In addition, when a given p-term of the Debye series contains a number of distinct contributions at the same scattering angle, the time-domain Fourier transform separates these as well (Fig. 14).

Figure 15 shows the results of time-domain calculations for the $p = 2$ term of the Debye series. The simultaneous relative minimum of the scattering angle at $\theta \approx 142°$ and the delay time at $t \approx 169$ fs correspond to the primary rainbow. Figure 15 also includes parametric curves showing the values of θ and t corresponding to $p = 2$ geometric rays with the impact parameter b obtained from

$$\Delta t = \Phi_{pl}/(kc), \tag{11.1}$$

where Φ_{pl} is the path-length-related phase in Eq. (5.33). Similarly, the path-length-related phase for $p = 0$ and $p = 1$ are given in Eqs. (12.31, 12.3).

Note the close agreement in Fig. 15 between the results for geometric rays with $0 \le b \le 1$ and the $p = 2$ Debye term. When $138° < \theta < 165.2°$, two supernumerary

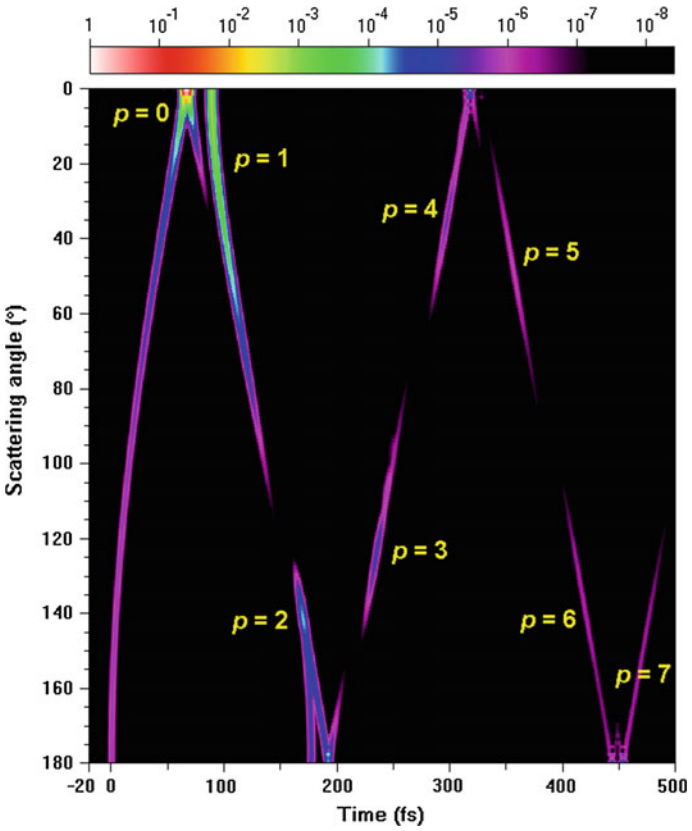


Fig. 14 Lorenz–Mie theory time-domain calculations for a spherical particle of radius $a = 10$ μm with $m_1 = 1.33257$ and $m_2 = 1$ for a 5 fs Gaussian pulse of red light (nominal wavelength $\lambda_0 = 0.65$ μm) as a function of scattering angle θ and the delay time t. The delay time is calculated with respect to that of the externally reflected pulse at $\theta = 180°$. The intensity of the scattered pulses is coded according to the false-color scale shown above the diagram. The delay time grid spacing is $\Delta t = 0.54$ fs and the scattering angle grid spacing is $\Delta\theta = 1°$

geometrical $p = 2$ rays are scattered at each value of θ (this was also illustrated in Fig. 13 for the specific case of $\theta = 150°$) and consequently generate distinct two pulses in the time-domain. Figure 15 shows that these pulses converge at the geometric rainbow angle of $\theta = 138°$ when $b = 0.8611$, and their separation increases for increasing θ. In the frequency-domain, interference between these two $p = 2$ geometrical rays produces a series of maxima and minima which are the supernumerary arcs of the primary rainbow, as can be seen from the $p = 2$ curves in

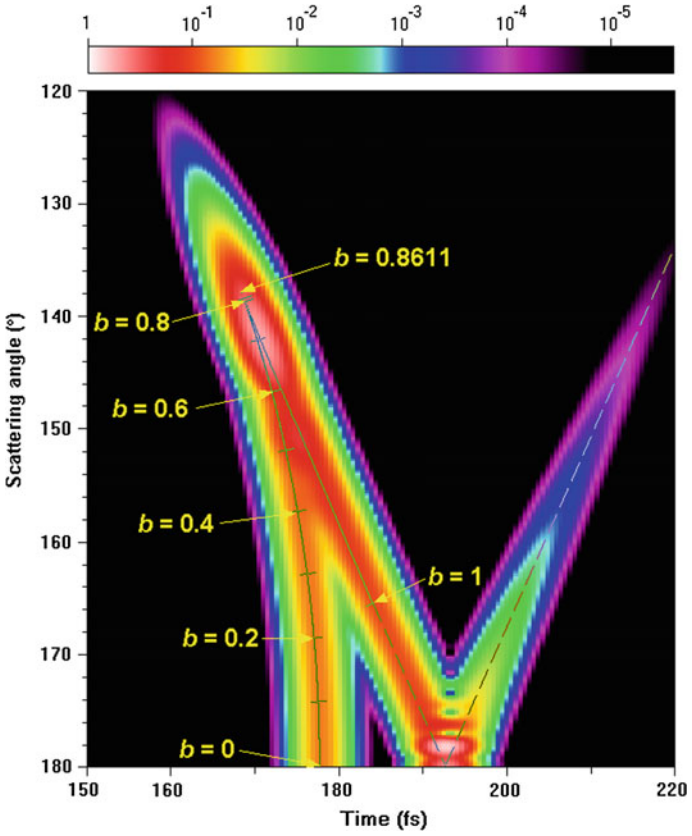


Fig. 15 As Fig. 14 but showing only the $p = 2$ Debye series term. The *solid parametric curves* plot the values of t and θ corresponding to $p = 2$ geometric rays for the specified values of impact parameter b where $0 \le b \le 1$. The dashed line extrapolating from the end of the *geometrical ray curve* shows the time delays corresponding to $p = 2$ surface waves. The delay time grid spacing is $\Delta t = 0.54$ fs and the scattering angle grid spacing is $\Delta\theta = 0.2°$

Fig. 3. When $165.2° < \theta < 180°$, only one $p = 2$ geometrical ray contributes to scattering at a given value of θ. Nevertheless, Fig. 15 shows two additional pulses with longer time delays in this angular interval. These are caused by two $p = 2$ surface waves, which will be discussed in greater detail in Sect. 12. These surface waves are also responsible for the atmospheric glory.

Figure 16 shows the time-domain results for the Debye series $p = 3$ term, which causes the second-order rainbow with its maximum intensity at $\theta \approx 125°$ and $t \approx 235$ fs. Again, the distinctive pattern in the time-domain is determined by two

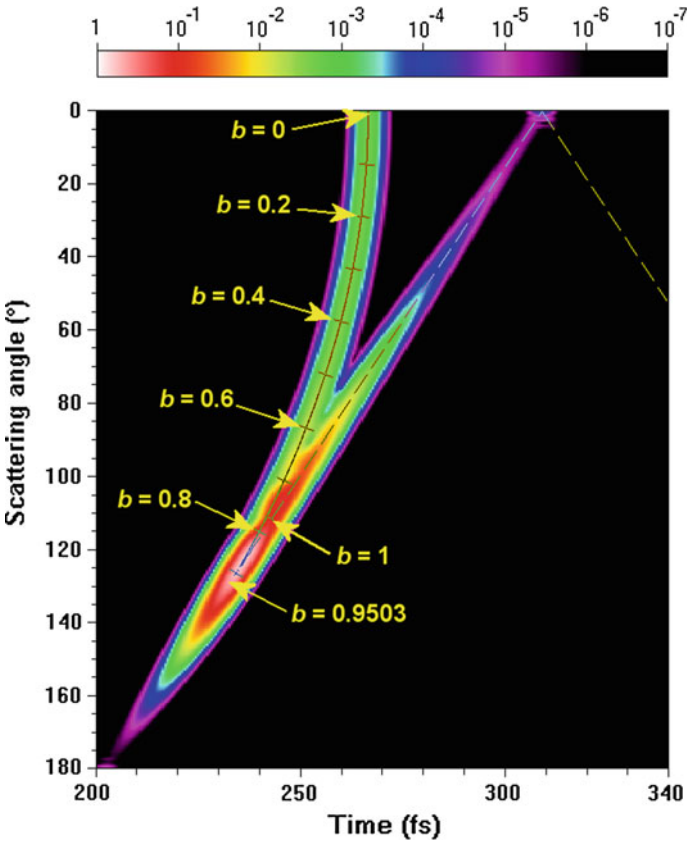


Fig. 16 As Fig. 15, but showing only the $p = 3$ Debye series term

geometrical $p = 3$ rays which converge at $\theta \approx 131°$ when $b = 0.9503$, supplemented by $p = 3$ surface waves that cross $\theta = 0°$ at $t = 266.5$ fs. In this case, the surface waves produce a forward scattering glory, unlike the $p = 2$ glory which is centered on $\theta = 180°$.

Figure 17 shows a set of time-domain results for scattering from coated spheres, illustrating how the various scattering mechanisms change in importance as the size of the core is increased. In this case, it is obvious that our understanding of scattering from coated spheres is dramatically enhanced by the time-domain approach.

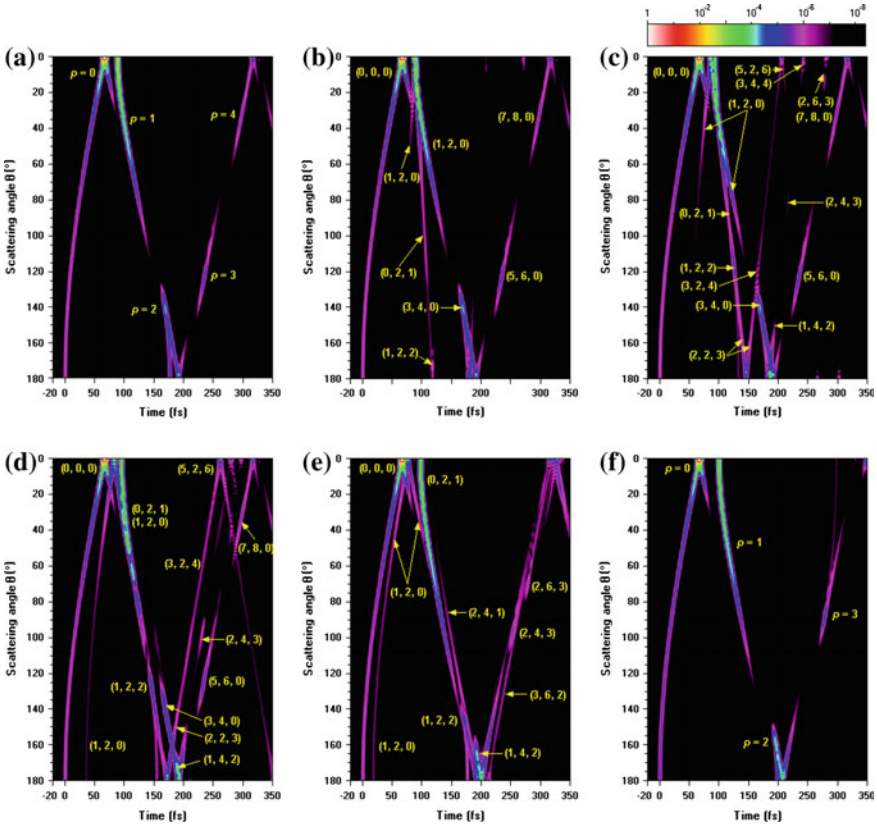


Fig. 17 Scattered intensity as a function of θ and t for a coated sphere of overall radius $a_{23} = 10$ μm where the core radius a_{12} is **a** 0, **b** 2, **c** 4, **d** 6, **e** 8, and **f** 10 μm. The refractive index of the core $m_1 = 1.5$, the refractive index of the coating $m_2 = 1.3333$, and the refractive index of the medium $m_3 = 1$. Diagrams **a** and **f** represent homogeneous spheres of refractive index $m = 1.3333$ and $m = 1.5$, respectively. The scattering angle grid spacing is $\Delta\theta = 1°$, and the three integers labeling each feature were described in Sect. 7

12 The Signature of Subtle Scattering Effects in Time-Domain Scattering

12.1 Electromagnetic Surface Waves

For each term of the Debye series, certain angular intervals contain the contribution of geometrical rays while other angular intervals contain no ray contributions. Whenever there is a transition between two such intervals, electromagnetic surface waves shed radiation into the zero-ray interval in order to smooth the transition. As a specific example, we discuss $p = 1$ scattering, i.e., transmission through the sphere of refractive index m when the external medium has unit refractive index.

For $ka \equiv x \gg 1$ when ray theory is valid, the transition between the one-ray scattering angle interval and the zero-ray interval occurs at the far-zone scattering angle

$$\theta^{\mathrm{C}} = 2\arccos(1/m). \tag{12.1}$$

In the angular interval $0^{\circ} < \theta < \theta^{\mathrm{C}}$, an approximation to the $p = 1$ far-zone scattered electric field for an x-polarized monochromatic plane wave of field strength E_0 incident on the sphere is derived from the sum over partial waves using the procedure of Sect. 5. In particular, the asymptotic approximation of Eqs. (5.20a, 5.20b) is used for the Lorenz–Mie angular functions $\pi_n(\theta)$ and $\tau_n(\theta)$. The geometrical ray contribution corresponding to incident rays striking the sphere surface (but not with grazing incidence) sums the partial waves contributions below the edge region $1 \le n \le x - \varepsilon_{\max} x^{1/3}$ with $\varepsilon_{\max} \approx 4.3$. The result is

$$\mathbf{E}^{\mathrm{ray}}_{p=1}(r,\theta,\varphi) = [\mathrm{i}\,E_0 \exp(\mathrm{i}kr)/(kr)]\left[S_2(\theta)\cos(\varphi)\mathbf{u}_{\boldsymbol{\theta}} - S_1(\theta)\sin(\varphi)\mathbf{u}_{\varphi}\right], \tag{12.2}$$

where

$$S_j(\theta) = \left[2\mathrm{i}x m^2 K_j/\left(m^2-1\right)^2\right]\exp\left\{2\mathrm{i}x + 2\mathrm{i}x\left[m^2 - 2m\,\cos(\theta/2) + 1\right]^{1/2}\right\}$$
$$\times\,[m\,\cos(\theta/2) - 1]^{3/2}[m - \cos(\theta/2)]^{3/2}/\left\{[\cos(\theta/2)]^{1/2}[m^2 - 2m\,\cos(\theta/2) + 1]\right\}. \tag{12.3}$$

The TE polarization is denoted by $j = 1$, the TM polarization is $j = 2$, and

$$\begin{array}{ll} K_j = 1 & \text{for } j = 1 \\ = \cos^2(\theta/2) & \text{for } j = 2. \end{array} \tag{12.4}$$

The ray theory scattering amplitude scales as the sphere size parameter x, and the ray scattered electric field vanishes for $\theta > \theta^{\mathrm{C}}$.

For $p = 1$ scattering in the interval $\theta^{\mathrm{C}} \le \theta \le 180^{\circ}$, the principal contribution to the electric field comes from partial waves in the lower half of the edge region $x - \varepsilon_{\max} x^{1/3} \le n \le x$ corresponding to rays striking the sphere surface with grazing incidence, plus partial waves in the upper half of the edge region $x < n \le x + \varepsilon_{\max} x^{1/3}$ corresponding to rays that classically just miss striking the sphere surface but interact with the centrifugal barrier surrounding it. In the edge region the large-argument, large-order approximation to the Riccati–Bessel, and Riccati–Neumann functions used in the derivation is proportional to the first and second Airy functions, $\mathrm{Ai}(z)$ and $\mathrm{Bi}(z)$, respectively [see Eq. (9.3.4) of Abramowitz and Stegun (1964)]. The simplest approximation to the partial wave sum in the edge region consists of converting it to an integral over ε, extending the limits of integration from $\pm\varepsilon_{\max}$ to $\pm\infty$, and recognizing the result as the modified Fock function (Nussenzveig 1965, 1969a, b)

$$f(s) = (1/2\pi)\exp(\mathrm{i}\pi/6) \int_{-\infty}^{\infty} \mathrm{d}w \exp(\mathrm{i}sw)/\{\mathrm{Ai}[w\exp(2\pi\mathrm{i}/3)]\}^2. \tag{12.5}$$

This procedure gives

$$S_j(\theta) \approx \left\{2x/\left[\left(m^2-1\right)\pi\sin(\theta)\right]\right\}^{1/2} \mathrm{i}L_j \exp\left\{\mathrm{i}\left[2x+2x\left(m^2-1\right)^{1/2}\right]\right\} \times \left\{\exp[\mathrm{i}(\pi/4+x\xi)]f\left[(x/2)^{1/3}\xi\right] + \exp[\mathrm{i}(-\pi/4+x\zeta)]f\left[(x/2)^{1/3}\zeta\right]\right\} \tag{12.6}$$

where

$$\begin{array}{ll} L_j = 1 & \text{for } j = 1 \\ = m^2 & \text{for } j = 2. \end{array} \tag{12.7}$$

The angle the two surface wave paths of Eq. (12.6) travel on the sphere surface is

$$\xi = \theta - \theta^{\mathrm{C}} \tag{12.8}$$

$$\zeta = 2\pi - \theta - \theta^{\mathrm{C}}. \tag{12.9}$$

The modified Fock function representation of $S_j(\theta)$ is a wave scattering phenomenon that smoothly interpolates between the far-zone geometrical optics angular region $0^\circ \le \theta < \theta^C$ and the far-zone surface wave-dominated angular region $\theta^{\mathrm{C}} < \theta < 180^\circ$. It is roughly analogous to the Fresnel region for light diffracted through an aperture that smoothly interpolates between the near-zone shadow-casting region and the far-zone Fraunhofer diffraction region.

The modified Fock function of Eq. (12.5) may be thought of mathematically as the Fourier transform of the integrand's denominator. For $s < -2$ in the geometrical optics limit of Eq. (12.5), the integral is most easily approximated using the method of stationary phase. For $s > 1$ in the surface wave-dominated limit, it is most easily evaluated by converting it to a contour integral and evaluating it by the residue method, including all the poles within the contour, which for convenience are numbered here by the integer $q \ge 1$. For $-2 < s < 1$ in the transition between the geometrical optics and surface wave-dominated limits, the integral is most easily evaluated numerically by rotating the $-\infty < w \le 0$ portion of the contour to the more rapidly converging path $z = r\exp(2\pi\mathrm{i}/3)$ for $0 \le r < \infty$.

In the surface wave-dominated angular region and considering only the residue of the strongest pole of the contour integral, $q = 1$, one has

$$S_j(\theta) \approx \left\{2^{1/6}x^{5/6}/\left[\left(m^2-1\right)\pi\sin(\theta)\right]^{1/2}\right\}\left[\mathrm{Ai}'(-X_1)\right]^{-2}L_j \times \left\{\xi\exp[\mathrm{i}\Phi_A(\theta)]\exp\left[-3^{1/2}x^{1/3}X_1\xi/2^{4/3}-L_j/\left(m^2-1\right)^{1/2}\right]\right. \left.+\,\zeta\exp[\mathrm{i}\Phi_B(\theta)]\exp\left[-3^{1/2}x^{1/3}X_1\zeta/2^{4/3}-L_j/\left(m^2-1\right)^{1/2}\right]\right\}, \quad (12.10)$$

where

$$\Phi_A(\theta) = \exp\left\{\mathrm{i}\left[2x+2x\left(m^2-1\right)^{1/2}+x\xi+x^{1/3}X_1\xi/2^{4/3}+7\pi/12\right]\right\} \quad (12.11a)$$

$$\Phi_B(\theta) = \exp\left\{\mathrm{i}\left[2x+2x\left(m^2-1\right)^{1/2}+x\zeta+x^{1/3}X_1\zeta/2^{4/3}+7\pi/12\right]\right\}. \quad (12.11b)$$

In addition, X_q are the zeros of the Airy function, $Ai(-X_q) = 0$, with $X_1 = 2.3381$ [see Table 10.13 of Abramowitz and Stegun (1964)]. Equation (12.10) requires a number of comments. Surface waves are not quite as strong as geometric ray contributions, scaling as $x^{5/6}$ rather than as x.

The second line of Eq. (12.10) describes the $q = 1$ clockwise-propagating surface wave of Fig. 18a where the length of its path along the sphere surface is $a\xi = a\ (\theta - \theta^{\mathrm{C}})$. As it propagates along the sphere surface, it continually tangentially sheds electromagnetic waves that propagate to the far zone. As these waves are shed, the strength of the remaining surface wave is exponentially damped. The amplitude factor of the TM surface wave is somewhat larger than that of the TE surface wave due to the L_j factor. But its damping rate is slightly faster, again due to the L_j factor. The third line of Eq. (12.10) describes the $q = 1$ counter-clockwise-propagating surface wave of Fig. 18b, where its path along the sphere surface is $a\zeta = a\ (360^{\circ} - \theta - \theta^{\mathrm{C}})$. These waves are angularly damped at the same rate as those in Fig. 18a. For $\theta^{\mathrm{C}} < \theta \ll 180^{\circ}$ they correspond to a longer path on the sphere surface and thus yield a smaller scattered field. As a result, this second surface wave is occasionally neglected with respect to the surface wave of Fig. 18a. If the residues of the $q \geq 2$ poles of the integrand of the modified Fock function are included in the residue evaluation of the contour integral, the value of X_q for these poles progressively increases, and these surface waves damp out at a progressively faster rate. The contributions of these additional poles are also occasionally neglected. Exact treatments of the scattering amplitude (Nussenzveig and Wiscombe 1987, 1991; Nussenzveig 1988) also lead to surface waves that circumnavigate the sphere surface any number of times before tangentially shedding their radiation. These are also occasionally neglected as well.

The frequency-domain plot of Fig. 19 for $x = 96.66$ and $m = 1.33257$ shows $|S_1(\theta)|^2$ and $|S_2(\theta)|^2$ for the $p = 1$ Debye series term. Also shown are the ray theory version of $|S_1(\theta)|^2$ and $|S_2(\theta)|^2$ and the surface wave version of Eq. (12.10)

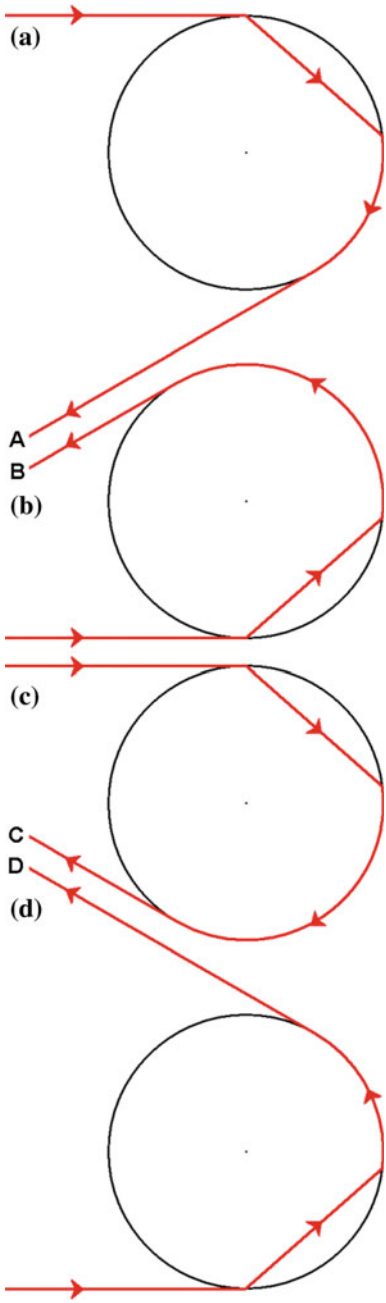


Fig. 18 Pictorial representation of the $p = 1$ surface waves A and B with $\theta < 180^{\circ}$, and C and D with $\theta > 180^{\circ}$. A and D are the short path length surface waves, and B and C are the long path length surface waves

considering only the residue of the $q = 1$ pole. One can easily check the accuracy of the magnitude of the surface wave contribution in this figure. However, the figure is not able to verify the accuracy of the phase of Eqs. (12.11a, 12.11b). This can be

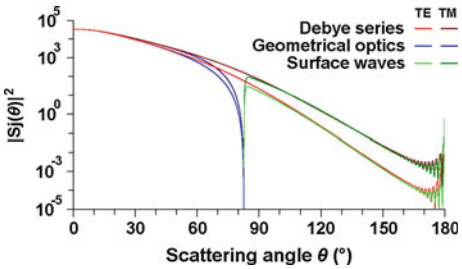


Fig. 19 TE and TM intensity of the $p = 1$ Debye series term as a function of angle for scattering of an x-polarized plane wave by a sphere with size parameter $x = 96.66$ and $m = 1.33257$. The ray contribution of Eqs. (12.13, 12.14) and the dominant surface wave contribution of Eqs. (12.10, 12.11a, 12.11b) are also shown

accomplished in the time-domain, however, since the velocity of a monochromatic surface wave along the curved path of the sphere surface is the slope of the time-domain trajectory in the surface wave region,

$$v/c = x/(\mathrm{d}\Phi/\mathrm{d}\theta) = \left[1 + X_1/\left(2^{4/3}x^{2/3}\right)\right]^{-1}. \tag{12.12}$$

Thus a scattered pulse should reach the detector slightly later than would be the case if $v = c$. As the size of the sphere increases, the approximations leading to Eq. (12.12) become more accurate while the predicted deviation of the wave velocity from c decreases. Conversely, as the size of the sphere becomes smaller, the analysis leading to Eq. (12.12) becomes less accurate while the predicted deviation of the wave speed from c increases. For $\lambda = 0.65$ μm and $a = 10$ μm, Eq. (12.12) predicts $v/c = 0.958$. When a pulse containing many different wavelengths is incident on the sphere, some component surface waves should travel slower, while others should travel faster. These two tendencies, however, do not completely cancel each other when the pulse spectrum is integrated. Rather, if the central wavelength in the pulse spectrum is λ_0, the effective surface wave speed of the pulse becomes

$$v/c \approx \left[1 + (1/3)X_1/\left(2^{4/3}x^{2/3}\right)\right]^{-1}. \tag{12.13}$$

For $\lambda_0 = 0.65$ μm and $a = 10$ μm, Eq. (12.13) predicts that $v/c = 0.9855$, which was found to be in good agreement with an analysis of Fig. 20 (Lock and Laven 2011a) that yielded $v/c = 0.9840$. A similar analysis of the $p = 2$ time-domain trajectory for the same pulse and sphere yielded $v/c = 0.9850$. These results illustrate that time-domain scattering is sensitive to both the amplitude and the phase of the scattered surface wave electric field.

Equation (12.12) assumed that surface waves propagate exactly along the sphere surface. van de Hulst has suggested an alternative interpretation in which surface

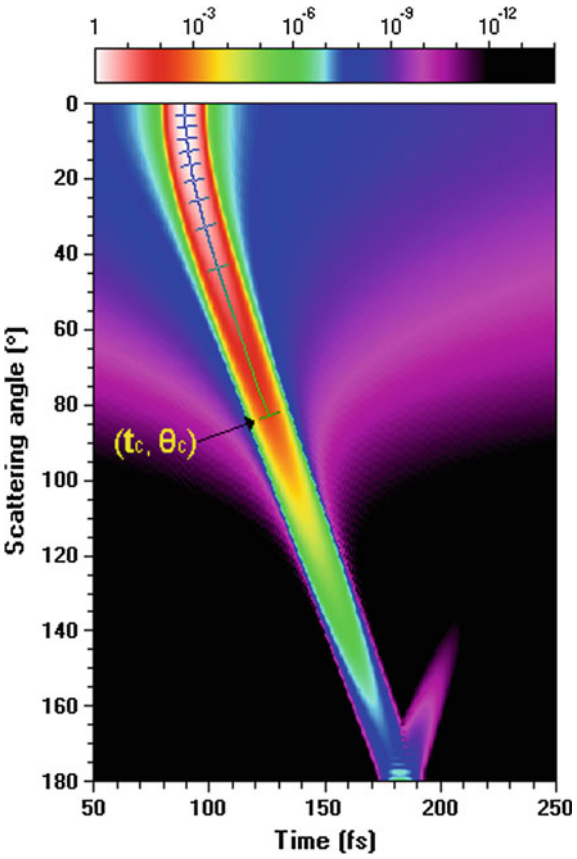


Fig. 20 $p = 1$ component of the scattered intensity in the time-domain as a function of the delay time t and scattering angle θ for the pulse and particle parameters of Fig. 14. The false-color intensity in Fig. 14 has been reset in Fig. 20 to emphasize the low-intensity components. The results of ray theory have been overlaid on the figure. The grid ticks on the ray theory line correspond to the ray impact parameter b in intervals of $\Delta b = 0.1$

waves for monochromatic plane wave incidence propagate with $v = c$, but do so slightly above the sphere surface [see p. 368 of van de Hulst (1981)], at

$$r \approx a\left[1 + X_1/\left(2^{4/3}x^{2/3}\right)\right] = a\left[1 + 0.9279/x^{2/3}\right]. \tag{12.14}$$

This interpretation is motivated as follows. For the partial wave $n + 1/2 = x$ at the center of the edge region, the scattered field contains, among other things, the Riccati–Bessel function

$$\psi_n(kr) = (\pi kr/2)^{1/2} J_{n+1/2}(kr) = (\pi kr/2)^{1/2} J_x(kr). \tag{12.15}$$

We parameterize the radial distance r outside the sphere in the vicinity of a by

$$r = x + zx^{1/3}, \tag{12.16}$$

so that

$$J_x(kr) \approx (2/x)^{1/3}\mathrm{Ai}\left(-2^{1/3}z\right). \tag{12.17}$$

The relative maximum of the principal peak of the Airy function occurs at $2^{1/3}\ z \approx 1.0188$ [see Table 10.13 of Abramowitz and Stegun (1964)], which gives

$$r \approx a\left(1 + 0.8086/x^{2/3}\right). \tag{12.18}$$

The value of r given by Eq. (12.14) occurs on the principal peak of the Airy function just past the relative maximum, consistent with the van de Hulst conjecture.

The situation becomes somewhat more complicated in the glory region

$$\theta = 180^\circ - \delta \tag{12.19}$$

for small δ where the clockwise-propagating and counter-clockwise-propagating surface waves, as shown in Fig. 18a, b, for $\delta > 0$ (i.e., $\theta < 180^\circ$), and in Fig. 18c, d for $\delta < 0$ (i.e., $\theta > 180^\circ$), are of comparable strength. In this angular regime, the Lorenz–Mie angular functions $\pi_n(\theta)$ and $\tau_n(\theta)$ are now approximated by the first term in their large-order asymptotic expansion for $\theta \approx 180^\circ$ [see p. 253 of van de Hulst (1981)]. The $p = 1$ far-zone scattered electric field in the glory region is then found to be

$$\begin{aligned}\mathbf{E}_{p=1}(r,\theta,\varphi) \approx\ & E_0[\exp(\mathrm{i}kr)/(kr)]G\left(\theta^C\right) \\ & \times \left\{\left(m^2-1\right)J_0(x\delta)\mathbf{u}_\mathbf{x} - \left(m^2+1\right)J_2(x\delta)\left[\cos(2\varphi)\mathbf{u}_\mathbf{x} + \sin(2\varphi)\mathbf{u}_\mathbf{y}\right]\right\},\end{aligned} \tag{12.20}$$

where

$$G\left(\theta^\mathrm{C}\right) \approx \left[x/\left(m^2-1\right)^{1/2}\right] f\left[(x/2)^{1/3}\left(\pi - \theta^\mathrm{C}\right)\right] \exp\left\{\mathrm{i}\left[2x + 2x\left(m^2-1\right)^{1/2} + x\left(\pi - \theta^\mathrm{C}\right)\right]\right\}. \tag{12.21}$$

In Eq. (12.21) both the modified Fock function and the phase have been approximated by their values at $\theta = 180^\circ$. The Bessel functions J_0 and J_2 in Eq. (12.20) describe the interference produced by axial focusing of the surface waves in all azimuthal planes in the near-backward direction. The θ-dependence of this focusing effect has been retained in Eq. (12.20). Evaluating the modified Fock function using the residue method, keeping only the contribution of the $q = 1$ pole, and using an average damping rate of the TE and TM surface waves, one obtains

$$G(\theta^{C}) \approx \left\{x^{4/3}(\pi-\theta^{C})/\left[2^{1/3}(m^2-1)^{1/2}\right]\right\}[\mathrm{Ai}'(-X_1)]^{-2} \times \exp\left\{-(\pi-\theta^{C})\left[3^{1/2}x^{1/3}X_1/2^{4/3}-(m^2+1)/2(m^2-1)^{1/2}\right]\right\} \times \exp\left\{\mathrm{i}\left[2x+2x(m^2-1)^{1/2}+x(\pi-\theta^{C})+x^{1/3}X_1(\pi-\theta^{C})/2^{4/3}-\mathrm{i}\pi/6\right]\right\}. \quad (12.22)$$

This glory focusing effect is stronger than geometrical optics, scaling as $x^{4/3}$ rather than as x. It should also be parenthetically noted that the atmospheric glory observed in scattering from water droplets in clouds is dominated by the analogous $p = 2$ surface wave glory (Nussenzveig 1969b; van de Hulst 1947; Bryant and Cox 1966; Bryant and Jarmie 1974; Laven 2005), also scaling as $x^{4/3}$.

The near-back-scattering enhancement evident in Fig. 19 is the $p = 1$ surface wave glory of an x-polarized monochromatic plane wave incident on a sphere, and is due to the interference of the four surface waves of Fig. 18a–d. The $p = 1$ time-domain trajectory in the glory region for a 5 fs plane wave pulse incident on the same sphere shown in Fig. 21a partially separates the contributions of the four surface waves. This figure contains more information about the individual surface waves than Fig. 19. For example, there is a set of intensity minima along the centerline of the X-shape structure in Fig. 21b at $\delta = \pm 1.0^{\circ}$, $\pm 3.0^{\circ}$, $\pm 5.0^{\circ}$, $\pm 6.8^{\circ}$, $\pm 8.7^{\circ}$, $\pm 10.5^{\circ}$, and $\pm 12.3^{\circ}$. These closely correspond to the interference minima of the shorter path surface waves of Fig. 18a or d with the longer path surface waves of Fig. 18b or c, respectively, which are -90° out of phase with respect to the shorter path contributions. The pulse version of Eq. (12.22) predicts that these minima should occur at $\delta = \pm 1.40^{\circ}$, $\pm 3.26^{\circ}$, $\pm 5.12^{\circ}$, $\pm 6.98^{\circ}$, $\pm 8.85^{\circ}$, $\pm 10.71^{\circ}$, and $\pm 12.57^{\circ}$ (Lock and Laven 2011a). In ray theory, the -90° phase shift is caused by the longer path surface waves crossing the back-scattering focal line [see p. 207 of van de Hulst (1981)]. In wave theory this shift does not occur discontinuously, but takes place over the angular width of the principal glory peak centered at $\theta = 180^{\circ}$.

12.2 The Time-Domain Signature of Diffraction

The analysis of the $p = 0$ term of the Debye series, diffraction-plus-external reflection, is more complicated than $p = 1$ because it consists of a number of different sub-processes that evolve and change from the far-zone near-forward scattering angle interval $\theta < (2/x)^{1/3}$, through the so-called transition region $(2/x)^{1/3} < \theta < M(2/x)^{1/3}$, to the large-angle interval $M(2/x)^{1/3} < \theta$, where $M \gg 1$, (e.g., $M = 6$ is often sufficient). In the first part of the near-forward interval, $0 \le \theta < 1/x$, the Lorenz–Mie angular functions $\pi_n(\theta)$ and $\tau_n(\theta)$ are proportional to combinations of the Bessel functions $J_0[(n + 1/2)\theta]$ and $J_2[(n + 1/2)\theta]$ (see Eqs. 5.7a, 5.7b) which is characteristic of the focusing of light scattered in the forward direction in all azimuthal planes. In the second part of the near-forward interval, $1/x < \theta < (2/x)^{1/3}$, the

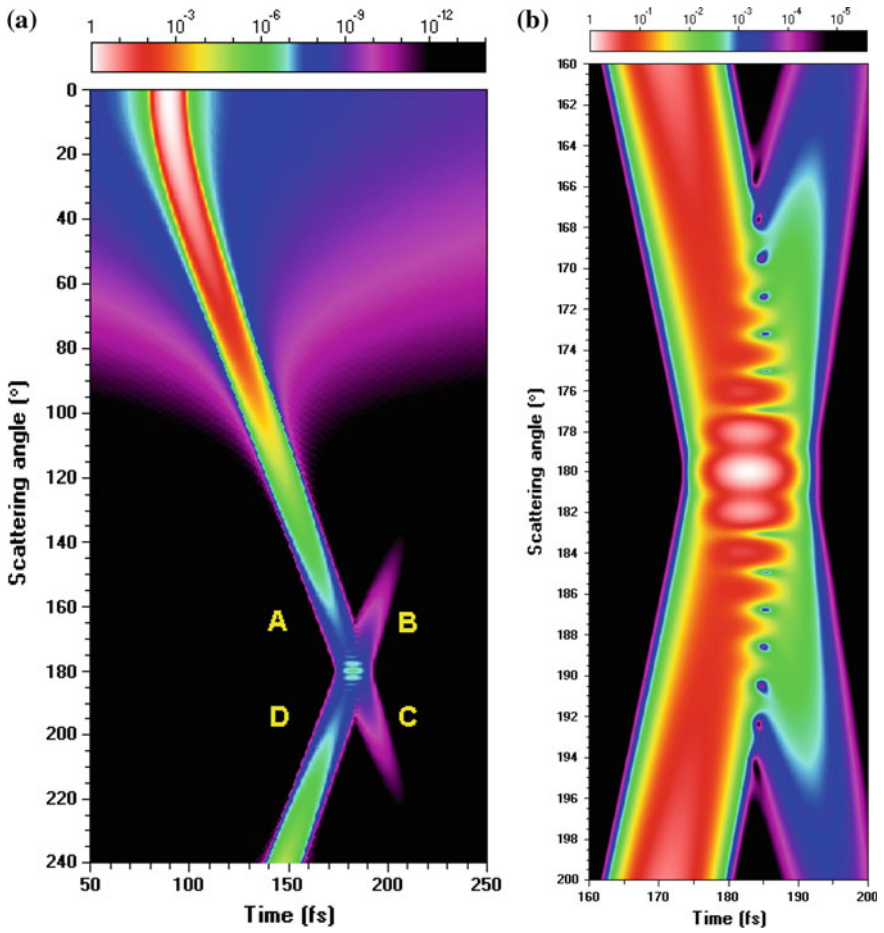


Fig. 21 **a** Similar to Fig. 20, but also showing scattering for $\theta > 180^{\circ}$. The surface waves $A + C$ and $D + B$ of Fig. 18 intersect at $\theta = 180^{\circ}$ to form the $p = 1$ surface wave glory. **b** Fine-resolution view of **a** in the vicinity of the $p = 1$ surface wave glory

Bessel functions J_0 and J_2 may be replaced by the first term of their asymptotic expansion (see Eqs. 5.20a, 5.20b). The three $p = 0$ sub-processes in the-near-forward angular interval are: (i) external reflection (denoted here by *er*) which consists of the $(1/2)(-R_n^{212})$ term summed over the partial waves $1 \le n \le x - \varepsilon_{\max} x^{1/3}$, (ii) diffraction (denoted by *d*) which consists of the (1/2) term summed over $1 \le n \le x$, and (iii) grazing-plus-tunneling reflection (denoted by *gr* + *tr*) which consists of the $(1/2)(-R_n^{212})$ term summed over the partial waves $x - \varepsilon_{\max} x^{1/3} < n \le x + \varepsilon_{\max} x^{1/3}$ plus the (1/2) term summed over $x < n \le x + \varepsilon_{\max} x^{1/3}$. These sub-processes are shown in the partial wave/amplitude diagram of Fig. 22. In the near-forward angular region the $p = 0$ Debye series amplitude may then be written as

$$S_j^{p=0}(\theta) = S_j^d(\theta) + S_j^{\mathrm{gr+tr}}(\theta) + S_j^{\mathrm{er}}(\theta). \tag{12.23}$$

For $0 \le \theta < 4/x$ the diffraction amplitude $S_j^d(\theta)$ is closely approximated by the usual Fraunhofer result

$$S_j^d(\theta) = x^2[J_1(x\theta)/(x\theta)], \tag{12.24}$$

that scales as x^2 due to forward focusing, and whose first zero occurs at

$$\theta \approx (1.22\pi)/x. \tag{12.25}$$

Farther out in the near-forward interval $4/x < \theta < (2/x)^{1/3}$, the asymptotic expansion of $J_1(x\theta)$ gives

$$S_j^d(\theta) \approx \{x/[2\pi \sin(\theta)]\}^{1/2}(1/\theta)[\exp(\mathrm{i}x\,\theta - 3\pi\mathrm{i}/4) + \exp(-\,\mathrm{i}x\,\theta + 3\pi\mathrm{i}/4)], \tag{12.26}$$

which scales as $x^{1/2}$, thus being weaker than ray scattering. The first few minima of the classical diffraction amplitude of Eq. (12.24) fall within the near-forward angular interval.

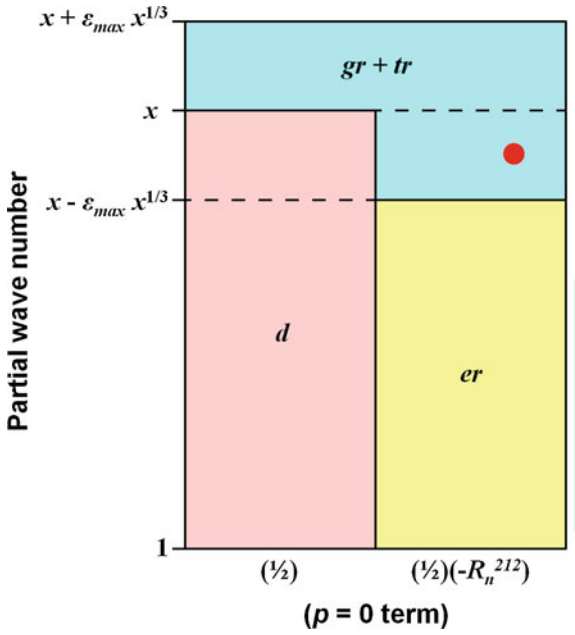


Fig. 22 $p = 0$ sub-processes of diffraction (*d*), grazing-plus-tunneling reflection (gr + tr), and geometrical ray external reflection (er) in the near-forward scattering angle interval. The stationary point of the gr + tr amplitude is shown as the *red dot*

The leading term of the grazing-plus-tunneling reflection amplitude $S_j^{\mathrm{gr+tr}}(\theta)$ in the near-forward angular interval is proportional to $x^{4/3}\, J_0(x\theta)$ which scales as a forward glory due to focusing of light in all azimuthal planes. The next largest term has contributions proportional to $x\, J_0(x\theta)$ and $x\, J_2(x\theta)$ which scale as ray scattering terms (Nussenzveig 1969a; Lock and Laven 2011b). This is because the stationary point of the external reflection portion of the partial wave sum occurs in the lower portion of the edge region of Fig. 22. The start and end of the transition region mentioned above are defined with respect to the location of the external reflection stationary point. The region corresponds to the partial wave number of the stationary point moving from somewhat above $n = x - \varepsilon_{\max}\, x^{1/3}$ in the lower portion of the edge region as shown in Fig. 22, to somewhat below it in the upper portion of the region of smaller partial waves as shown in Fig. 23. When the stationary point occurs at the partial wave $n = x - \varepsilon_{\max}\, x^{1/3}$ in the middle of the transition region, ray optics gives the scattering angle as

$$\theta = 2^{7/6}\varepsilon_{\max}^{1/2}(2/x)^{1/3} = 4.6(2/x)^{1/3}. \tag{12.27}$$

Lastly, in the near-forward angular region $S_j^{\mathrm{er}}(\theta)$ of Eq. (12.23) is small and may be ignored because the terms of the external reflection partial wave sum in the region of smaller partial waves in Fig. 22 are slowly varying in amplitude but rapidly varying in phase. The surviving combination of diffraction and grazing-plus-tunneling reflection for near-forward scattering is pictorially illustrated in Fig. 24.

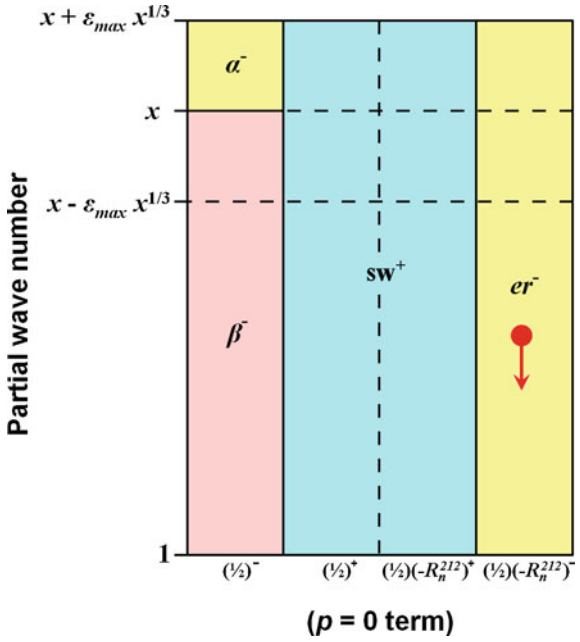


Fig. 23 $p = 0$ sub-processes of surface waves (sw^+), geometrical ray external reflection (er^-), and the processes α^- and β^- which cancel each other in the large scattering angle interval. The stationary point of the er^- amplitude is shown as the *red dot*

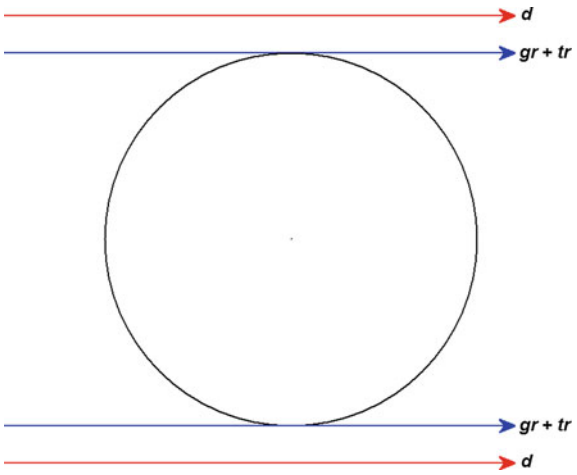


Fig. 24 Pictorial illustration of the $p = 0$ sub-processes of diffraction and grazing-plus-tunneling reflection in the near-forward scattering angle interval

Turning now to the large scattering angle interval $\theta > M\,(2/x)^{1/3}$, the dominant Lorenz–Mie angular function $\tau_n(\theta)$ of Eq. (5.20b) is proportional to

$$\cos[(n+1/2)\theta - \pi/4] = (1/2)\exp[\mathrm{i}(n+1/2)\theta - \mathrm{i}\pi/4] + (1/2)\exp[-\mathrm{i}(n+1/2)\theta + \mathrm{i}\pi/4]. \tag{12.28}$$

This leads to sub-processes whose scattering amplitudes $S_j^+(\theta)$ and $S_j^-(\theta)$ include either the first or second term of the right-hand side of Eq. (12.28), respectively. These new sub-processes are shown in the partial wave /amplitude diagram of Fig. 23. The $p = 0$ scattering amplitude in the large-angle interval then consists of four sub-processes, external reflection (denoted here by er^-), $p = 0$ surface waves (denoted by sw^+), as well as the contributions of the partial wave intervals α^- and β^-. These last two contributions turn out to be equal and opposite due to Babinet's principle, and cancel each other. The $p = 0$ scattering amplitude in the large-angle interval may then be written as

$$S_j^{p=0}(\theta) = S_j^{\mathrm{er}-}(\theta) + S_j^{\mathrm{sw}+}(\theta). \tag{12.29}$$

The external reflection sum over partial waves $S_j^{\mathrm{er}-}(\theta)$ is dominated by the contribution of the stationary point at

$$(n+1/2)_{\mathrm{sp}} = x\cos(\theta/2), \tag{12.30}$$

which migrates deeper into the region of smaller partial waves, er^-, as the scattering angle increases. This is noted in Fig. 23. In this regime one obtains the geometrical ray scattering result,

$$S_j^{\mathrm{er}-}(\theta) = \mathrm{i}(x/2) r_j(\theta) \exp[-2\mathrm{i}x \sin(\theta/2)], \tag{12.31}$$

where the Fresnel external reflection coefficient is

$$r_j(\theta) = \left\{ \left[m^2 - \cos^2(\theta/2) \right]^{1/2} - L_j \sin(\theta/2) \right\} / \left\{ \left[m^2 - \cos^2(\theta/2) \right]^{1/2} + L_j \sin(\theta/2) \right\}. \tag{12.32}$$

The ray theory scattering amplitude of Eq. (12.31) scales as x. The surface wave amplitude $S_j^{\mathrm{sw}+}(\theta)$ again scales as $x^{5/6}$ and has the same exponential falloff as was the case for $p = 1$ surface waves (Nussenzveig 1969a). The external reflection and surface wave contributions to the $p = 0$ scattering amplitude are pictorially illustrated in Fig. 25.

The angular interval $(2/x)^{1/3} < \theta < M\ (2/x)^{1/3}$ marks the gradual transition via Fock functions and modified Fock functions from the diffraction and grazing-plus-tunneling reflection amplitudes for smaller angles to the geometrical ray and surface wave amplitudes for large angles (Nussenzveig and Wiscombe 1987, 1991; Nussenzveig 1988). Said in a more pictorial way, it corresponds to the gradual splitting of the (1/2) and (1/2) $(-R_n^{212})$ amplitudes of Fig. 22 into $(1/2)^{\pm}$ and (1/2) $(-R_n^{212})^{\pm}$ amplitudes of Fig. 23, accompanied by a gradual joining together of the partial waves in the edge region with those in the region of smaller partial waves, and the gradual joining together of the amplitudes $(1/2)^{+}$ and $(1/2)(-R_n^{212})^{+}$ into a single amplitude which gives rise to electromagnetic surface waves.

The rearrangement of the partial wave sums in the transition region is not evident in the frequency-domain in Fig. 26 for $\lambda = 0.65$ μm, $a = 10$ μm, and $m = 1.33257$. It is difficult to distinguish the different $p = 0$ sub-processes and to see where the transition region begins and ends. On the other hand, much of this structure becomes separated in the time-domain. The $p = 0$ time-domain trajectory for a 5 fs

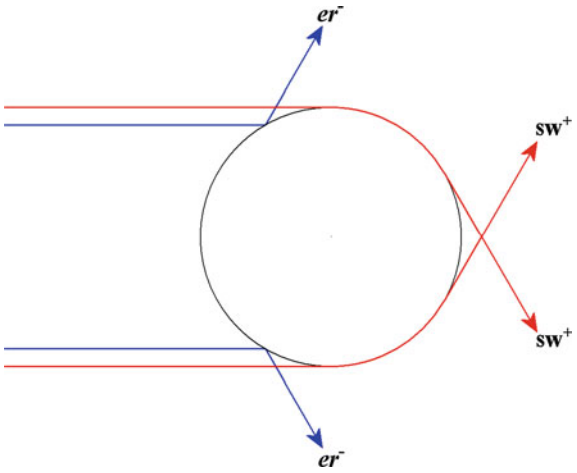


Fig. 25 Pictorial illustration of the $p = 0$ sub-processes of geometrical ray external reflection and surface waves in the large scattering angle interval

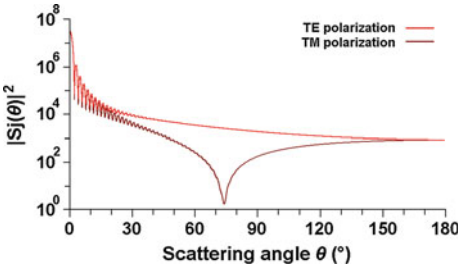


Fig. 26 $p = 0$ Debye series term of the TE and TM intensity as a function of angle for scattering of an x-polarized plane wave by a sphere with $\lambda = 0.65$ μm, $a = 10$ μm and $m = 1.33257$

Gaussian plane wave pulse with $\lambda_0 = 0.65$ μm incident on a sphere with $a = 10$ μm and $m = 1.33257$ is shown in Fig. 27a, and the near-forward angular interval is shown in greater detail in Fig. 27b. For these parameters, the transition region is predicted to begin at $\theta \approx 15.73^\circ$ and its midpoint, according to Eq. (12.27), should be at $\theta \approx 72.36^\circ$. Diffraction which scales as x^2 and grazing-plus-tunneling reflection which scales as a combination of $x^{4/3}$ and x should completely overlap in the first part of the near-forward angular interval. According to Eq. (12.26), they should then split at $\theta \approx 2.37^\circ$ into an inverted-V structure whose straight diagonal arms in the second part of the near-forward interval have the slope

$$d\theta/dt = \pm c/a. \tag{12.33}$$

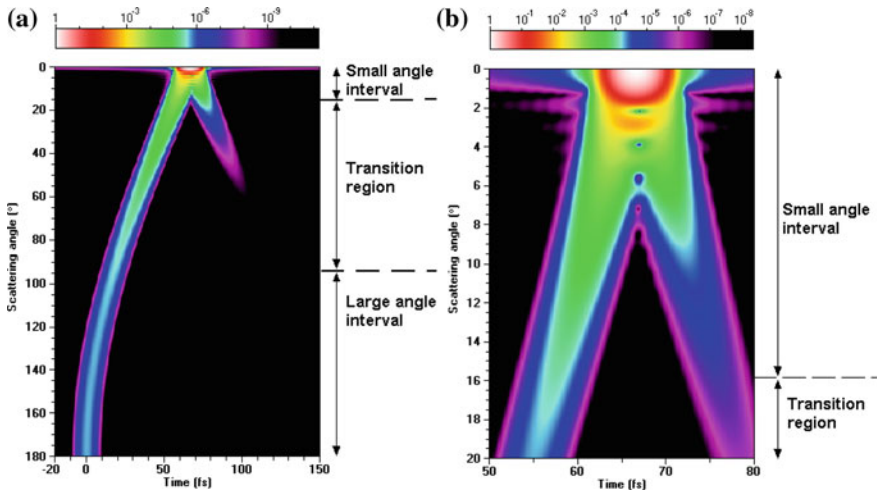


Fig. 27 **a** Intensity of the $p = 0$ Mie-Debye term as a function of the scattering angle and delay time for the pulse and particle parameters of Fig. 14. **b** Fine-resolution view of **a** in the near-forward direction. The fundamental sampling interval is $\Delta t = 0.135$ fs

These predictions agree well with Fig. 27b. In the transition region, the smaller time-delay arm of diffraction is gradually replaced by external reflection, and the longer time-delay arm is gradually replaced by $p = 0$ surface waves. The details of this gradual replacement are evident in Fig. 6 of Lock and Laven (2011b).

13 Summary

It is often said that everything that happens, happens for a reason. In the context of light scattering by a single homogeneous sphere, "everything that happens" is implicit in the formulas of Lorenz–Mie theory since they are, after all, the exact solution to the posed electromagnetic boundary value problem. It is one thing to know the exact solution, but it is quite another to understand all the implications hiding within it. The Debye series with its decomposition of the partial wave scattering amplitudes into diffraction of the partial waves, external reflection, and transmission following any number of internal reflections goes a long way toward explaining "the reason" why the scattered intensity has the structure it does, when considered in the short wavelength limit.

But even within a single Debye series term, it often happens that a number of physical sub-processes angularly overlap and interfere. When this happens, time-domain scattering takes up where the Debye series left off and successfully separates these overlapping and interfering sub-processes. In doing do, it opens up a new level of meaning concerning the details of electromagnetic scattering in particular, and of the nature of the interaction of micron-sized matter with radiation in general. One can best and most cleverly use various light scattering effects as a tool for technological advancement only when one firmly understands the physical nature and implications of the phenomenon. The exposition given in this chapter makes the claim that the Debye series and time-domain scattering are our two most fruitful tools for achieving this understanding.

References

Abramowitz M, Stegun IA (eds) (1964) Handbook of mathematical functions. National Bureau of Standards, Gaithersburg

Aden AL, Kerker M (1951) Scattering of electromagnetic waves from two concentric spheres. J Appl Phys 22:1242–1246

Adler CL, Lock JA, Stone BR, Garcia CJ (1997) High-order interior caustics produced in scattering of a diagonally incident plane wave by a circular cylinder. J Opt Soc Am A 14:1305–1315

Arfken G (1985) Mathematical methods for physicists, 3rd edn. Academic

Asano S, Yamamoto G (1975) Light scattering by a spheroidal particle. Appl Opt 14:29–49 (1975); errata: Appl Opt 15:2028 (1976)

Bakić S, Heinisch C, Damaschke N, Tschudi T, Tropea C (2008) Time integrated detection of femtosecond laser pulses scattered by small droplets. Appl Opt 47:523–530

Bech H, Leder A (2004) Particle sizing by ultrashort laser pulses—numerical simulation. Optik 115:205–217
Bech H, Leder A (2006) Particle sizing by time-resolved Mie calculations—a numerical study. Optik 117:40–47
Berry MV, Mount KE (1972) Semiclassical approximations in wave mechanics. Rep Prog Phys 35:315–397
Berry MV, Upstill C (1980) Catastrophe optics: morphologies of caustics and their diffraction patterns. Prog Opt 18:257–346
Berry MV, Nye JF, Wright FJ (1979) The elliptic umbilic diffraction catastrophe. Phil Trans Roy Soc Lon 291:453–484
Bhandari R (1985) Scattering coefficients for a multilayered sphere; analytic expressions and algorithms. Appl Opt 24:1960–1967
Bohren CF, Huffman DR (1983) Absorption and scattering of light by small particles. Wiley, Hoboken
Born M, Wolf E (1980) Principles of optics, 6th edn. Cambridge University Press, Cambridge
Bryant HC, Cox AJ (1966) Mie theory and the glory. J Opt Soc Am 56:1529–1532
Bryant HC, Jarmie N (1974) The glory. Sci Am 231:60–71
Chen YM (1964) Diffraction by a smooth transparent object. J Math Phys 5:820–832
Chew WC (1990) Waves and fields in inhomogeneous media. Van Nostrand
Debye P (1908) Das Elektromagnetische Feld um einen Zylinder und die Theorie des Regenbogens. Phys Zeit 9:775–778, reprinted and translated into English (1994) In: Marston PL (ed) Geometrical aspects of scattering, MS89. SPIE Press, pp 198–204
Felsen LB, Marcuvitz N (1973) Radiation and scattering of waves. Prentice-Hall, Upper Saddle River
Flammer C (1957) Spheroidal wave functions. Stanford University Press, Palo Alto
Ford KW, Wheeler JA (1959) Semiclassical description of scattering. Ann Phys NY 7:259–286
Franz W (1954) Uber die Greenschen Funktionen des Zylinders und der Kugel. Zeit F Natur 9a:705–716
Glantschnig WJ, Chen S-H (1981) Light scattering from water droplets in the geometrical optics approximation. Appl Opt 20:2499–2509
Gouesbet G (2003) Debye series formulation for generalized Lorenz–Mie theory with the Bromwich method. Part Part Syst Charact 20:382–386
Gouesbet G, Maheu B, Gréhan G (1988) Light scattering from a sphere arbitrarily located in a Gaussian beam, using a Bromwich formulation. J Opt Soc Am A 5:1427–1443
Griffiths DJ (2005) Introduction to quantum mechanics, 2nd edn. Pearson Prentice-Hall, Upper Saddle River
Han YP, Méès L, Ren KF, Gréhan G, Wu ZS, Gouesbet G (2004) Far scattered field from a spheroid under a femtosecond pulsed illumination in a generalized Lorenz–Mie theory framework. Opt Commun 231:71–77
Hecht E (1987) Optics, second edition. Addison-Wesley, Boston
Hovenac EA, Lock JA (1992) Assessing the contributions of surface waves and complex rays to far-field Mie scattering by use of the Debye series. J Opt Soc Am A 9:781–795
Johnson BR (1996) Light scattering by a multilayer sphere. Appl Opt 35:3286–3296
Kai L, Massoli P (1994) Scattering of electromagnetic plane waves by radially-inhomogeneous spheres: a finely-stratified sphere model. Appl Opt 33:501–511
Kerker M (1969) The scattering of light and other electromagnetic radiation. Academic
Khare V (1976) Short wavelength scattering of electromagnetic waves by a homogeneous dielectric sphere. PhD dissertation, University of Rochester, Rochester, NY
Laven P (2005) How glories are formed. Appl Opt 44:5675–5683
Laven P (2010) Separating diffraction from scattering: the million dollar challenge. J Nanophotonics 4:041593
Laven P (2011) Time-domain analysis of scattering by a water droplet. Appl Opt 50:F29–F38
Laven P, Lock JA (2012) Understanding light scattering by a coated sphere. Part 2: Time domain analysis. J Opt Soc Am A 29:1498–1507

Li R, Han X, Jiang H, Ren KF (2006a) Debye series for light scattering by a multilayered sphere. Appl Opt 45:1260–1270
Li R, Han X, Jiang H, Ren KF (2006b) Debye series of normally incident plane-wave scattering by an infinite multilayered cylinder. Appl Opt 45:6255–6262
Li R, Han X, Shi L, Ren KF, Jiang H (2007) Debye series for Gaussian beam scattering by a multilayer sphere. Appl Opt 46:4804–4812
Li R, Han X, Ren KF (2009) Generalized Debye series expansion of electromagnetic plane wave scattering by an infinite multilayered cylinder at oblique incidence. Phys Rev E 79:036602
Lock JA (1988) Cooperative effects among partial waves in Mie scattering. J Opt Soc Am A 5:2032–2044
Lock JA (1993) Contribution of high-order rainbows to the scattering of a Gaussian laser beam by a spherical particle. J Opt Soc Am A 10:693–706
Lock JA (1996) Ray scattering by an arbitrarily oriented spheroid. II. Transmission and cross-polarization effects. Appl Opt 35:515–531
Lock JA (2001) Semiclassical scattering of an electric dipole source inside a spherical particle. J Opt Soc Am A 18:3085–3097
Lock JA (2005) Debye series analysis of scattering of a plane wave by a spherical Bragg grating. Appl Opt 44:5594–5603
Lock JA (2008) Scattering of an electromagnetic plane wave by a Luneburg lens. II. Wave theory. J Opt Soc Am A 25:2980–2990
Lock JA (2011) Scattering of a transversely confined Neumann beam by a spherical particle. J Opt Soc Am A 28:2577–2587
Lock JA, Adler CL (1997) Debye-series analysis of the first order rainbow produced in scattering of a diagonally incident plane wave by a circular cylinder. J Opt Soc Am A 14:1316–1328
Lock JA, Hovenac EA (1991) Internal caustic structure of illuminated liquid droplets. J Opt Soc Am A 8:1541–1552
Lock JA, Laven P (2011a) Mie Scattering in the time domain. Part 1. The role of surface waves. J Opt Soc Am A 28:1086–1095
Lock JA, Laven P (2011b) Mie scattering in the time domain. Part II. The role of diffraction. J Opt Soc Am A 28:1096–1106
Lock JA, Laven P (2012) Understanding light scattering by a coated sphere. Part I. Theoretical considerations. J Opt Soc Am A 29:1489–1497
Lock JA, Jamison JM, Lin C-Y (1994) Rainbow scattering by a coated sphere. Appl Opt 33:4677–4690, 4960
Mackowski DW, Altenkirch RA, Menguc MP (1990) Internal absorption cross sections in a stratified sphere. Appl Opt 29:1551–1559
Méès L, Gouesbet G, Gréhan G (2001a) Scattering of laser pulses (plane wave and focused Gaussian beam) by spheres. Appl Opt 40:2546–2550
Méès L, Gréhan G, Gouesbet G (2001b) Time-resolved scattering diagrams for a sphere illuminated by a plane wave and focused short pulses. Opt Commun 194:59–65
Méès L, Gouesbet G, Gréhan G (2009) Transient internal and scattered fields from a multi-layered sphere illuminated by a pulsed laser. Opt Commun 282:4189–4193
Nussenzveig HM (1965) High-frequency scattering by an impenetrable sphere. Ann Phys (NY) 34:23–95
Nussenzveig HM (1969a) High-frequency scattering by a transparent sphere. I. Direct reflection and transmission. J Math Phys 10:82–124
Nussenzveig HM (1969b) High-frequency scattering by a transparent sphere. II. Theory of the rainbow and the glory. J Math Phys 10:125–176
Nussenzveig HM (1979) Complex angular momentum theory of the rainbow and the glory. J Opt Soc Am 69:1068–1079, 1193–1194
Nussenzveig HM (1988) Uniform approximation in scattering by spheres. J Phys A Math Gen 21:81–109
Nussenzveig HM (1992) Diffraction effects in semiclassical scattering. Cambridge University Press, Cambridge

Nussenzveig HM, Wiscombe WJ (1987) Diffraction as tunneling. Phys Rev Lett 59:1667–1670
Nussenzveig HM, Wiscombe WJ (1991) Complex angular momentum approximation to hard-core scattering. Phys Rev A 43:2093–2112
Rubinow SI (1961) Scattering from a penetrable sphere at short wavelengths. Ann Phys NY 14:305–332
Shen J, Wang H (2010) Calculation of Debye series expansion of light scattering. Appl Opt 49:2422–2428
Stratton JA (1941) Electromagnetic theory. McGraw-Hill
Tam WG, Corriveau R (1978) Scattering of electromagnetic beams by spherical objects. J Opt Soc Am 68:763–767
Twersky V (1964) On propagation in random media of discrete scatterers. In: Bellman R (ed) Stochastic processes in mathematical physics and engineering. Symp Appl Math XVI, Proc, pp 84–116
Ungut A, Gréhan G, Gouesbet G (1981) Comparisons between geometrical optics and Lorenz–Mie theory. Appl Opt 20:2911–2918
van de Hulst HC (1947) A theory of the anti-coronae. J Opt Soc Am 37:16–22
van de Hulst HC (1981) Light scattering by small particles. Dover
van der Pol B, Bremmer H (1937) The diffraction of electromagnetic waves from an electrical point source round a finitely conducting sphere, with applications to radiotelegraphy and the theory of the rainbow, part 2. Phil Mag 24:825–864
Wait JR (1955) Scattering of a plane wave from a circular dielectric cylinder at oblique incidence. Can J Phys 33:189–195
Walker JD (1976) Multiple rainbows from single drops of water and other liquids. Am J Phys 44:421–433
Wiscombe WJ (1980) Improved Mie scattering algorithms. Appl Opt 19:1505–1509
Wu Z-S, Li H-Y (2008) Debye series of scattering by a multi-layer cylinder in an off-axis 2D Gaussian beam. Chin Phys Lett 25:1672–1675
Wu ZS, Wang YP (1991) Electromagnetic scattering for multilayered sphere: recursive algorithms. Radio Sci 26:1393–1401
Wu ZS, Guo LX, Ren KF, Gouesbet G, Gréhan G (1997) Improved algorithm for electromagnetic scattering of plane waves and shaped beams by multilayered spheres. Appl Opt 36:5188–5198
Xu F, Lock JA (2010) Debye series for light scattering by a coated nonspherical particle. Phys Rev A 81:063812
Xu F, Lock JA, Tropea C (2010a) Debye series for light scattering by a spheroid. J Opt Soc Am A 27:671–686
Xu F, Lock JA, Gouesbet G (2010b) Debye series for light scattering by a nonspherical particle. Phys Rev A 81:043824

Morphological Models for Inhomogeneous Particles: Light Scattering by Aerosols, Cometary Dust, and Living Cells

Michael Kahnert, Timo Nousiainen and Johannes Markkanen

1 Introduction

Particle populations in nature can often be composed of different chemical components that are either externally or internally mixed. By "external mixture", we refer to an ensemble of particles in which different chemical species are contained in physically separated particles; an "internal mixture" refers to the case in which different chemical components are contained in the same particles. Examples of internal mixtures are liquid-phase aerosols containing water and dissolved sodium chloride. In that case, the two chemical species are *homogeneously* mixed on the molecular level. Other examples are solid-phase light-absorbing carbon (LAC) aggregates onto which a coating of liquid-phase material has condensed. In the latter case, we are dealing with an *inhomogeneous* mixture of different chemical species. It is such inhomogeneous particles that will be the focus of this chapter.

Inhomogeneous particles are ubiquitous in the terrestrial atmosphere, in the solar system, and in the interstellar medium. For instance, mineral dust aerosols are often

M. Kahnert (✉)
Swedish Meteorological and Hydrological Institute, Folkborgsvägen 17,
SE-601 76 Norrköping, Sweden
e-mail: michael.kahnert@smhi.se

T. Nousiainen
Finnish Meteorological Institute, Erik Palménin aukio 1, FI-00560 Helsinki, Finland
e-mail: timo.nousiainen@fmi.fi

J. Markkanen
Department of Physics, University of Helsinki, Gustaf Hällströmin katu 2a,
FI-00560 Helsinki, Finland
e-mail: johannes.markkanen@helsinki.fi

© Springer-Verlag Berlin Heidelberg 2016

A. Kokhanovsky (ed.), *Light Scattering Reviews, Volume 11*,
Springer Praxis Books, DOI 10.1007/978-3-662-49538-4_6

composed of different mineral components, such as carbonates of calcium or magnesium, or oxides containing silicon, aluminium, iron, sodium, potassium, or calcium (e.g. Lindqvist et al. 2014). Anthropogenic aerosols can consist of inhomogeneous mixtures of LAC, organic compounds, sulphate, nitrate, ammonium, and water (e.g. McFiggans et al. 2006). Also, materials of anthropogenic origin, such as sulphates or nitrates, can be inhomogeneously mixed with natural aerosols, such as mineral dust or sea salt.

Desert dust aerosols result from wind re-suspension of soil dust which, in turn, originates from eolic erosion processes of rocks. Thus, the inhomogeneous mineralogical composition of mineral dust particles reflects that of the parent rock. By contrast, the main causes of internal mixing in anthropogenic terrestrial aerosols are dynamic processes, such as condensation and coagulation. For instance, as LAC aggregates age in the atmosphere they are oxidised and, as a result, become more hydrophilic and more compact (e.g. Coz and Leck 2011). Subsequent condensation processes produce LAC aerosols that can be covered with liquid phase materials (e.g. Bond et al. 2006).

Cosmic dust particles are often the result of dynamic growth processes quite similar to those of terrestrial aerosols, although operating on different time scales. For instance, in the outflow of red-giant stars refractory elements, such as silicon, oxygen, and carbon, nucleate to form solid particles of silicon oxide or silicon carbide. Such particles may later coagulate to form aggregates. Farther away from the star, more volatile elements, such as carbon monoxide, water, carbon dioxide, molecular oxygen and nitrogen, ammonia, or organic compounds can condense onto the particles and form icy outer layers (e.g. Whittet 2003).

Volcanic ash particles are generated during explosive volcanic eruptions. They are composed of crustal material with glassy or crystalline structure and different degrees of vesicularity. The vesicles often form in high-viscosity magmas from expanding gas contained in the magma. The resulting particles have irregular shapes and inhomogeneous composition characterised by gas vesicles entrained inside the solidified crustal particles.

Modelling the optical properties of particles with irregular shapes, complex morphology and heterogeneous chemical composition is a formidable task. Different approaches ranging from effective-medium approximations (EMAs) and simplified regular-shape models to morphologically realistic quasi-replicas of real-world particles have been applied in the past. The level of sophistication in the choice of a model particle is often determined by the intended use of the model, and it is constrained by the availability of computational resources, numerical limitations of electromagnetic scattering software and the incomplete knowledge of the morphology and composition of real particles. In environmental modelling applications, such as in climate models or in remote sensing retrieval systems, a priori information on morphological properties is often difficult to obtain. This lack of

information in conjunction with computational constraints often motivates the use of drastically simplified particle models. By contrast, more fundamental studies that aim at understanding the impact of particle morphology on radiometric properties must make use of more elaborate and more realistic particle models. Light-scattering studies based on realistic model particles can also be useful for assessing the errors in climate models or remote-sensing retrieval algorithms that are introduced by simplified particle-shape models.

In general, the very concept of a *model* implies certain simplifications, based on making a discrimination between important and less relevant factors that contribute to a certain phenomenon (e.g. Eddington 1926). The use of exact replicas of real particles does not follow such an approach. Replica particles are most useful for producing results that are as realistic as possible. Following Kahnert et al. (2014), one may therefore refer to such particles as "numerical laboratories" rather than "model particles". Gaining insight into the physics of the electromagnetic scattering process is often easier when using model particles with a degree of complexity that lies in between the two extremes of over-simplified geometries and real-world replicas. In particular, it can be advantageous to use model particles that selectively focus on one specific morphological aspect, while neglecting other complicating factors. Such general aspects of the choice of model particles are discussed in more detail in Kahnert et al. (2014). In the following sections, we will discuss several examples that illustrate the use of model particles and numerical laboratories and explain the advantages of different approaches with regard to the intended use of the models. We start with a short discussion of the merits and limitations of effective-medium approximations in Sect. 2. In Sect. 3, we cover LAC aerosols that are inhomogeneously mixed with sulphate, organic material, salt, or water. Section 4 is devoted to mineral dust aerosols. This is a field of extensive research; however, the discussion in this section will focus only on those issues that pertain to the inhomogeneity of mineral particles. Section 5 deals with volcanic ash particles with entrained gas vesicles. Section 6 deals with cometary dust particles. Finally, Sect. 7 gives a brief introduction to modelling optical properties of living cells.

Readers who want to obtain more information are referred to the recent monograph by Babenko (2010). An introduction to light scattering media optics in general, which contains several pedagogical examples of inhomogeneous particles, solution methods, and approximations, can be found in the book by Kokhanovsky (2004).

2 Effective-Medium Approximations

The main part of this review will deal with the construction of model particles that explicitly account for inhomogeneities in the dielectric properties of particles. But before we enter this subject, we discuss an alternative and, arguably, simpler

approach that is based on determining some average dielectric properties of heterogeneous materials. The goal is to specify an effective refractive index, m_{eff}, in such a way that the heterogeneous particle has the same optical properties as the corresponding homogeneous particle with a refractive index m_{eff}. The specification of m_{eff} is based on the refractive indices of each of the components in the heterogeneous mixture. These methods are known as EMAs. Such an approach can be motivated by a desire to reduce numerical complexities, or by our lack of knowledge of the detailed morphology of inhomogeneous particles, which can make any attempts to explicitly account for inhomogeneities rather unpromising. Here, we will not attempt to review the various EMA approaches; we refer to Chýek et al. (2000), Sihvola (1999). Rather, we will discuss some aspects regarding the validity and limitations of EMAs.

In the literature, several examples can be found for cases in which EMA approaches yield valid results. For instance, it is often taken for granted that LAC is a homogeneous material that can be described by a single refractive index. In reality, LAC contains a variable amount of void fraction, where the "void inclusions" are on the nanometre scale. By use of EMA, one can model m_{eff} as a function of the void fraction. A plot of the imaginary versus the real part of m_{eff} yields a so-called void-fraction curve. It has been shown in Bond and Bergstrom (2006) that many measurements of the refractive index of LAC reported in the literature lie along that void-fraction curve. This indicates that EMA is, indeed, a valid approach for describing the dielectric properties of LAC, and that a variation in the void fraction can often explain discrepancies among measurements of the refractive index of LAC.

One can also find modelling studies which support the validity of effective-medium approximations, at least under certain circumstances. For instance, Videen and Chýek (1998) compared optical efficiencies and differential scattering optical properties of inhomogeneous spheres to those of homogeneous spheres, where different approaches were tested for modelling m_{eff}. The spherical host particle was assumed to be water with $m_1 = 1.335$, and the spherical inclusion was LAC with $m_2 = 1.94 + 0.66i$. The radius of the composite particle was assumed to be equal to the wavelength of the incident radiation. It was found that the integral optical properties were reproduced to within 1 %, as long as the size parameter x_{inc} of the LAC inclusion was sufficiently small, i.e., $x_{\text{inc}} < 0.5$. In another study by Mishchenko et al. (2014), spherical hosts with multiple spherical inclusions were considered. The host and the inclusions had generic refractive indices of 1.33 and 1.55, respectively. The size parameter of the particle was fixed at $2\pi R/\lambda = 10$. The number of inclusions was varied up to $N = 8000$. It was found that the elements of the Stokes scattering matrix are well reproduced by EMA over the entire range of scattering angles, provided that the size parameter of the inclusions is less than about 0.5, and the number of inclusions is sufficiently large.

However, there are also counter-examples that call into question the validity of EMAs. For instance, in Kocifaj and Videen (2008) the optical properties of inhomogeneous particles composed of LAC, ammonium sulphate and organic matter were modelled. It was found that EMAs tend to underestimate the scattering efficiency and overestimate the asymmetry parameter and absorption efficiency. Similar observations were made in Kahnert et al. (2013), in which a model of LAC fractal aggregates encapsulated in a spherical shell of sulphate was considered.

These examples illustrate that one can find examples in the literature on both sides of the argument. When considering the evidence reported thus far one should be aware that most studies consider only a limited range of inhomogeneous morphologies, sizes or dielectric materials. This raises the question how such findings can be generalised. A related question is if it is valid to interpret measurements on inhomogeneous particles by assuming a homogeneous material. In doing so, one may even "retrieve" an effective refractive index m_{eff}. This question has been brought up in Bohren and Huffman (1983), Bohren (1986). However, it is not clear if such an approach would yield a useful optical constant that allows us to predict optical properties even for other experimental situations, such as for particles of different sizes or morphologies, or for measurements of other optical properties. One would like to find a way to specify an effective refractive index that correctly predicts the optical properties of inhomogeneous materials over the entire range of relevant parameters.

Thus, to test the validity of effective dielectric constants, it is essential to perform such tests over a large range of particle parameters. This is hampered by the high computational demands of such efforts. As a first step in this direction, a recent study by Kahnert (2015) has considered spheres with $m_1 = 1.6$ (representing a generic mineral material) with spherical hematite inclusions, $m_2 = 3.102 + 0.0925i$, sizes in the range of 80 nm–2.8 μm, hematite volume fractions of 1 and 4 % (representing low and high hematite contents in terrestrial desert dust aerosols), and three different morphologies, namely with 1, 10, and 100 randomly placed inclusions. Figure 1 shows an example for a sphere with 30 inclusions and a hematite

Fig. 1 Inhomogeneous mineral sphere with 30 randomly placed hematite inclusions. This model was employed in Kahnert (2015) to assess the validity of effective-medium approximations

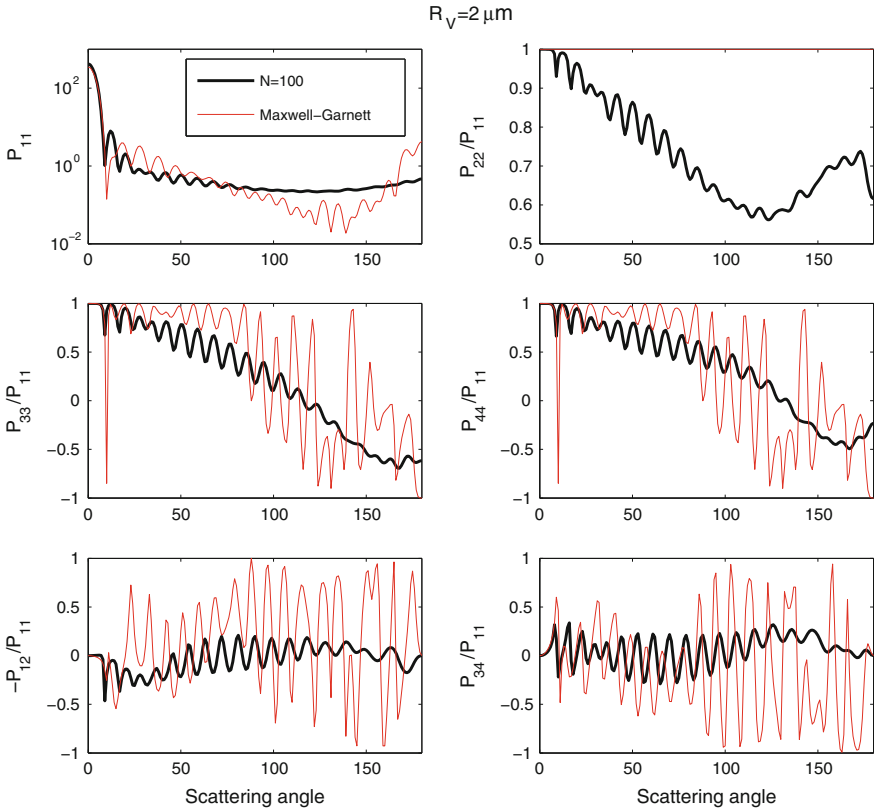


Fig. 2 Elements of the scattering matrix of a sphere with radius $R_V = 2$ µm and a hematite volume fraction of 4 % at a wavelength of 550 nm. An inhomogeneous sphere model with 100 spherical hematite inclusions (*black*) and a homogeneous sphere model based on EMA (*red*) are compared

volume fraction of 4 %. To assess the variability in the radiometric properties, ten stochastic realisations of the random geometry have been considered for each particle size, volume fraction and number of inclusions. The results have been compared to various EMAs. Also, it has been attempted to find a best-fit of the radiometric properties by freely varying the real and imaginary parts of m_{eff}. Although the best-fit approach did achieve some limited improvements over conventional EMAs, it turned out not to be possible to reproduce all radiometric properties with a homogeneous particle model over the entire range of parameters considered in that study. Also, it was found that the elements of the Stokes scattering matrix as well as the single-scattering albedo vary considerably among the particles of the same size and volume fraction, but with 1, 10, and 100 inclusions. As an example, Fig. 2 shows the elements of the scattering matrix of a mineral/hematite particle with a radius of 2 µm at a wavelength of 550 nm. The two

model particles that are being compared are an inhomogeneous sphere containing 100 spherical hematite inclusions with a volume fraction of 4 % (black line), and a corresponding homogeneous sphere with an effective refractive index obtained with the Maxwell Garnett EMA (red line). The scattering matrix of the inhomogeneous particles has been averaged over orientations and over 10 stochastic realisation of the inhomogeneous morphology. The limitations of the homogeneous sphere model are obvious. It underestimates the phase function at side-scattering angles, and it overestimates the phase function in the backscattering direction. Also, it severely overestimates the oscillations with scattering angle of the elements P_{33}, P_{44}, P_{12}, and P_{34}. Most strikingly, the element P_{22}, which enters into the definition of the depolarisation ratio, is as low as 0.6 for inhomogeneous particles, while it is identically equal to 1 for spherically symmetric particles. A closer analysis (not shown) revealed that the scattering matrix elements obtained with the two models agreed well only for small particles; they start to diverge already at a particle radius of about 300 nm. This corresponds to an inclusion size parameter of 0.25. Note that this is only half as large as the critical size parameter of 0.5 reported in Mishchenko et al. (2014). However, the optical contrast between the inclusions and the host medium was much smaller in Mishchenko et al. (2014) than it was in Kahnert (2015). This could indicate that the critical inclusion size parameter below which the inhomogeneous material is well described by a suitable EMA becomes smaller the higher the optical contrast between the mixed materials is.

The studies discussed above indicate that particle inhomogeneity may be a very significant morphological property, and that simple homogeneous particle models can be insufficiently accurate to reproduce the radiometric properties, unless the inclusion size parameter is sufficiently small. If it is not, then even if one should be able to find an m_{eff} that yields good agreement between the optical properties of inhomogeneous particles with corresponding EMA computations at some selected particle sizes, volume fractions or inhomogeneous morphologies, one cannot be confident that such agreements will extend outside the limited range of parameters for which the EMA approach has been validated. The studies by Mishchenko et al. (2014), Kahnert (2015) used simplified spherical hosts with spherical inclusions. The idea was to isolate the effect of particle inhomogeneity from other effects, such as overall nonsphericity, surface roughness or aggregation, which are often simultaneously present in natural particles. When working with more realistic numerical laboratories, it is often difficult to disentangle the impact of different morphological properties on radiometric quantities. On the other hand, homogeneous spheres have rather unique optical properties. A comparison of optical properties of homogeneous and inhomogeneous spheres, such as in Mishchenko et al. (2014), Kahnert (2015), may, therefore, overstate the significance of inhomogeneity. A recent study by Videen et al. (2015) on irregular agglomerated debris particles showed little differences among the orientation-averaged phase functions computed for particles with inhomogeneous and homogeneous compositions. Both an EMA and an external mixture approach had been tested for the homogeneous particles. Replacing the irregular by spherical particles caused errors that far exceeded those related to the homogeneous particle approximation. However, it

was also found that the degree of linear polarisation is more sensitive to inhomogeneous composition than the phase function. The study was limited to a single size parameter of 10.

It thus seems that particle shape and surface roughness may partially mask, or even dominate over the impact of inhomogeneity. However, more dedicated studies are required to test this hypothesis. Performing such tests for a sufficiently large parameter space of particle sizes and volume fractions will require formidable computational resources.

3 Encapsulated Light-Absorbing Carbon Aggregates

Light-absorbing carbon (LAC) aerosols are formed during incomplete combustion of fossil fuel, biofuel and biomass. Common emission sources are biomass burning, diesel vehicles, industrial processes and power generation. LAC consists of incompletely graphetised amorphous carbon that can be partially oxidised. As a result of condensation processes in the atmosphere, LAC aerosols become coated by chemical compounds such as sulphate (Adachi and Buseck 2008; Worringen et al. 2008), organic materials (Adachi and Buseck 2008), sea salt (Scarnato et al. 2013) or water. While the real and imaginary parts of the refractive index of LAC are both relatively high throughout the short-wave part of the spectrum (e.g. Chang and Charalampopoulos 1990; Bond and Bergstrom 2006), the coating material is typically weakly absorbing and optically much softer than LAC at UV and visible wavelengths (e.g. Hess et al. 1998; see also Fig. 4 in Kahnert 2010a). Thus, the dielectric properties of such particles are characterised by a high optical contrast between the LAC core and the liquid-phase shell.

Model geometries in light-scattering computations range from highly simplified shapes to morphologically realistic numerical laboratories, of which several are shown in Fig. 3. The simplest internal mixture model is the homogeneous sphere model in which the mixture of LAC and the coating material is described by EMA (Fig. 3d). Here, we will concentrate on model particles that explicitly account for the dielectric inhomogeneity of the particles.

Among the simplest inhomogeneous model particles are the concentric core-shell model (Ackerman and Toon 1981) (Fig. 3e), the non-concentric core-shell model (Fuller 1995) and dimers of concentric core-shell particles (Fuller et al. 1999). Studies based on non-concentric core shell models revealed that the optical properties of encapsulated geometries are mostly sensitive to the LAC volume fraction (Fuller et al. 1999); the positioning of the LAC core inside the spherical shell has little effect on the absorption cross section. This is, at least, true for particle sizes and LAC volume fractions that are typical for atmospheric aerosols (Bond et al. 2006). For particles with large coatings and correspondingly low LAC volume fractions, such as those encountered in LAC-containing cloud droplets, particles with concentrically positioned LAC cores can have significantly enhanced absorption cross sections compared to off-centre geometries (Fuller et al. 1999).

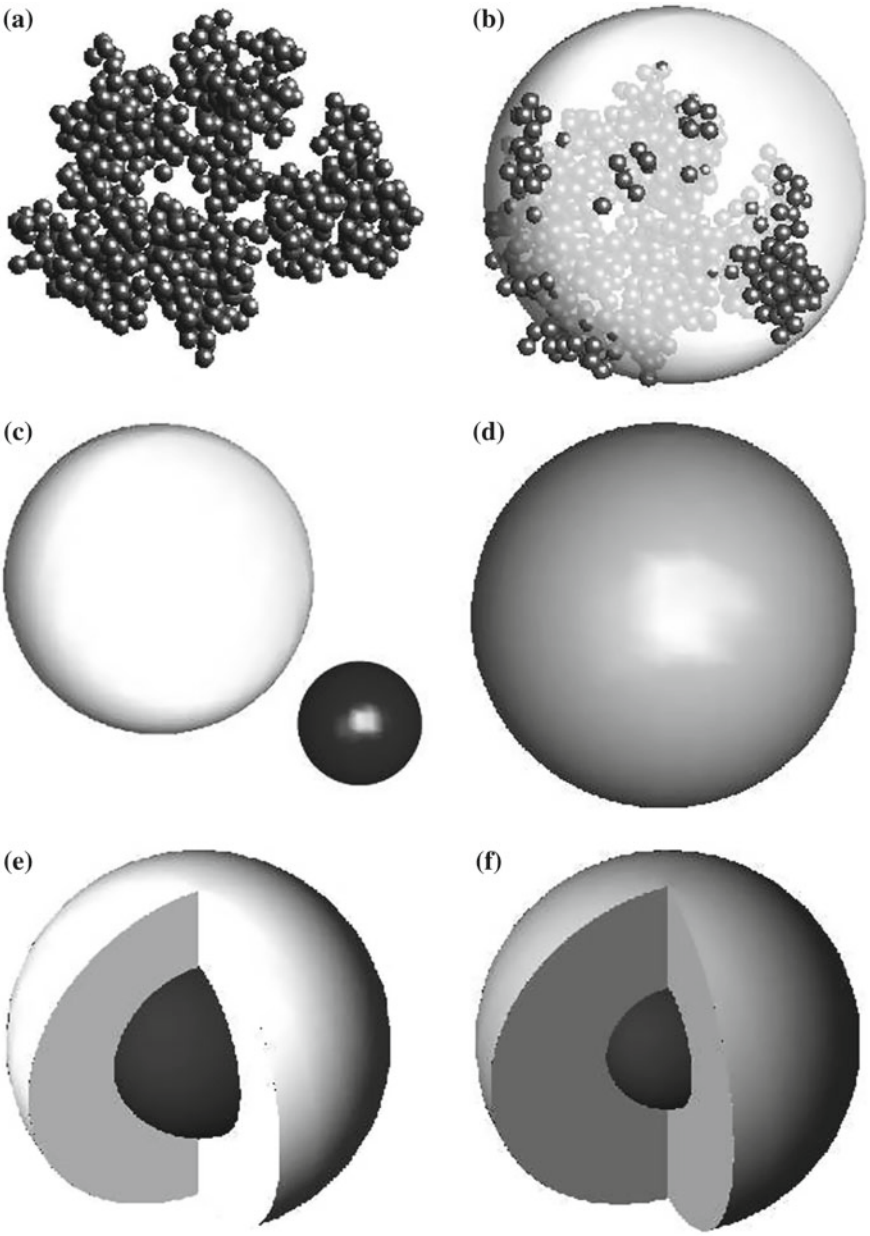


Fig. 3 Examples of LAC model particles. **a** Bare aggregate; **b** encapsulated aggregate; **c** externally mixed homogeneous spheres; **d** internally mixed homogeneous sphere/EMA; **e** core-shell; **f** core-grey-shell

Models with multiple spherical LAC inclusions inside a sulphate shell have been considered (Worringen et al. 2008), where each LAC particle was modelled by a size-equivalent homogeneous sphere. The computed optical cross sections, single-scattering albedo, and asymmetry parameter were compared to those of simpler model particles, such as core-shell geometries with an LAC core and a sulphate shell, core-shell particles with a sulphate core and an LAC shell, as well as hemispheric LAC/sulphate and sandwich sulphate/LAC/sulphate geometries. The differences among the various model particles were mostly dependent on the LAC volume fraction, and they varied among different optical parameters.

Laboratory and field observations by use of transmission electron microscopy (Adachi and Buseck 2008; Worringen et al. 2008) and electron tomography (van Poppel et al. 2005; Adachi et al. 2010) have provided detailed information on the morphology of encapsulated LAC particles. This has paved the way for going beyond simple coated sphere models. Most importantly, the observational data yield information on the complex fractal aggregate structure of the LAC cores. Before discussing inhomogeneous mixtures of LAC and liquid-phase aerosol components, a few remarks on the morphology of *bare* aggregates and their optical properties are in order (see Fig. 3a).

A fractal aggregate composed of N_s spherical monomers of radius a can be described by the statistical scaling relation (Jones 2006)

$$N_\mathrm{s} = k_0 \left(\frac{R_\mathrm{g}}{a} \right)^{D_\mathrm{f}}. \tag{1}$$

Here D_f and k_0 denote, respectively, the fractal dimension and structural prefactor of the aggregate, and

$$R_\mathrm{g} = \sqrt{\frac{1}{N_\mathrm{s}} \sum_{i=1}^{N_\mathrm{s}} r_i^2} \tag{2}$$

R_g denotes the radius of gyration, which is defined in terms of the distances r_i of the ith monomer from the aggregates's centre of mass. A low-fractal dimension corresponds to a more lacy aggregate; whereas, a high-fractal dimension corresponds to a more compact arrangement of the monomers.

The set of parameters $\{a, N_\mathrm{s}, D_\mathrm{f}, k_0\}$ describes a class of fractal aggregates that contains an infinite number of individual geometries. Thus, this approach allows us to classify a myriad of fractal geometries by a limited set of parameters. We may view these classes as "geometric equivalence classes". However, this classification would be useless for the purpose of optical modelling unless the optical properties show a very low variance among the members of each class. In other words, these geometric equivalence classes should, within a tolerable variance, also constitute "optical equivalence classes". In recent studies, optical properties were computed for several stochastic realisations of bare (Kahnert 2010b) and encapsulated LAC

aggregates (Kahnert et al. 2012) belonging to the same class of fractal aggregates, and the variation in the optical properties within a fractal aggregate class was investigated. It was found that the extinction, scattering and absorption cross section, and, related to that, the single-scattering albedo showed almost no variation among particles of the same geometry class. The variation in the backscattering cross sections and asymmetry parameter was slightly larger, but still within just a few percent. The linear and circular backscattering depolarisation ratio of the encapsulated geometries varied by up to a factor of 2–3 for particles larger than about 400 nm; (the study was performed at one UV, one visible, and one NIR wavelength). Thus for the integral optical properties, the geometric equivalence classes can, for all practical purposes, be considered optical equivalence classes. However, for differential optical properties as sensitive to morphology as depolarisation, the description of particle properties in terms of fractal parameters does not seem to provide a useful classification that constrains depolarisation within tolerable limits.

While many of the optical properties of fractal aggregates, bare or coated, do not vary significantly within each class of fractal aggregates, they can vary significantly among particles belonging to different classes. For instance, in Liu et al. (2008) a large range of the parameter space of bare aggregates was covered; D_f was varied between 1.25 and 3, N_s was varied between 200 and 800, and two values of the monomer radius, $a = 15$ nm and $a = 25$ nm, were considered. Also, two values were used for the refractive index of soot, $m = 1.75 + 0.5i$ and $m = 2 + i$. These ranges of D_f and m are quite large in comparison to those encountered in real atmospheric soot particles; however, these large ranges were chosen in order to brace the range of parameters relevant for realistic applications. In general, the optical cross sections, single-scattering albedo and asymmetry parameter were quite sensitive to a variation in D_f and N_s. Interestingly enough, the scattering cross section and the single-scattering albedo increased with D_f, which indicates the increasing importance of electromagnetic interaction among the monomers as the aggregate becomes more compact.

In Kahnert (2010c), the sensitivity of the mass absorption cross section, MAC, and the single-scattering albedo of bare aggregates to a variation in m, N_s, and a was investigated. The fractal dimension was fixed at $D_\mathrm{f} = 1.8$. The refractive index was varied within a range that was based on different measurements of m (Bond and Bergstrom 2006). This range was much more constrained than that in Liu et al. (2008). The monomer radius a was varied between 10 and 25 nm, which, again, was based on measurements (Bond and Bergstrom 2006). It was found that MAC was mainly sensitive to a variation in m, and rather insensitive to changes in a. On the other hand, a variation in a mainly impacted the single-scattering albedo. The largest values of $|m|$ gave the best agreement of computed MAC with observations, and a value of a around 25 nm yielded the best agreement of modelled single-scattering albedos with observations, where the experimental reference values were taken from the review in Bond and Bergstrom (2006). This agrees well with field observations (Adachi et al. 2010), where an average monomer radius of $a = 22$ nm was reported, with a standard deviation of 6 nm.

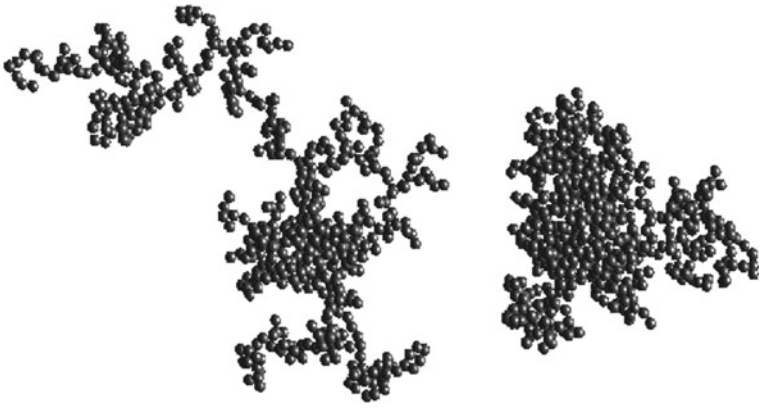


Fig. 4 Examples of fractal aggregates with fractal parameters $D_f = 1.82$, $k_0 = 1.27$ (*left*), and $D_f = 2.4$, $k_0 = 0.7$ (*right*)

Frequently used values of the fractal parameters are $D_f = 1.82$, $k_0 = 1.27$ (Bond and Bergstrom 2006), which corresponds to relatively lacy aggregates—see Fig. 4 (left). However, this value has been called into question by modern 3D electron tomography measurements (Adachi et al. 2007), which suggest that bare LAC aggregates in the atmosphere are often more compact with fractal parameters around $D_f = 2.4$, $k_0 = 0.7$—see Fig. 4 (right). It is quite possible that a value of $D_f = 1.82$, which is often observed in laboratory experiments, is typical for freshly produced LAC, while the value observed in the field of $D_f = 2.4$ is more representative of moderately aged atmospheric LAC aggregates. Based on this observational evidence, a recent study of the sensitivity of optical properties of bare LAC aggregates to a variation in D_f concentrated on the range from 1.8 to 2.4 (Kahnert and Devasthale 2011). The computed optical properties were coupled to a chemical transport model to generate several realistic scenarios of vertical aerosol profiles representing clean marine and polluted continental air masses over land or ocean surfaces. The vertical profiles of aerosol optical properties for these different scenarios were coupled to a radiative-transfer model to quantify the radiative-forcing effect of LAC. Corresponding computations were performed based on size-equivalent homogeneous LAC spheres. The results showed that the radiative-forcing effect computed with an aggregate model can be more than twice as high as that obtained with a homogeneous sphere model. Also, lacy aggregates give forcing estimates that are 10–60 % higher than those obtained with compact aggregates.

As LAC ages in the atmosphere, the aggregates oxidise and become more hydrophilic. Subsequent condensation of liquid-phase material can significantly alter the optical properties. For instance, experimental studies have shown that coating of LAC with secondary organic aerosols can increase the specific absorption cross section (Schnaiter et al. 2003). This increase can be as large as a factor of two relative to bare LAC aggregates (Schnaiter et al. 2005), which agrees with

model simulations (Scarnato et al. 2013). Condensation and evaporation of water vapour onto the particles contributes to a progressive restructuring (Colbeck et al. 1990; Ramachandran and Reist 1995), resulting in an increase in the fractal dimension, i.e. in a compactification of the aggregate structure. In field observations of LAC aggregates internally mixed with sulphate and organic matter, fractal parameters as high as $D_f = 2.6$ and $k_0 = 1.2$ were reported (Adachi and Buseck 2008), while a recent electron-tomography analysis of field samples reported a somewhat lower value of $D_f = 2.2$ with a standard deviation of 0.2 (Adachi et al. 2010). The same field observations yielded an LAC volume fraction of 0.07 ± 0.08, and a relative LAC position of 0.54 ± 0.35. The relative position is defined as d/R, where d is the distance between the centres of mass of the LAC aggregate and the encapsulating host particle, and R is the volume-equivalent radius of the entire particle (LAC + host).

Largely based on these observational data, recent modelling studies (Kahnert et al. 2012, 2013) have employed model geometries consisting of LAC aggregates with $D_f = 2.6$, $k_0 = 1.2$, $a = 25$ nm, and N_s in the range of 4–1600 (as an example, see Fig. 3b). The aggregates were encapsulated by a spherical shell of sulphate, where the total particle size was in the range $R = 100$–500 nm, and the position of the aggregate was $d/R = 0.5$. Two volume fractions were considered, $f = 0.07$ (as a typical case) and $f = 0.20$ (as a more extreme case). The optical properties were computed for three selected wavelengths in Kahnert et al. (2012), and for the full spectral range from 200 nm to 12.2 μm in Kahnert et al. (2013). The computations were repeated for several simplified particle geometries; a comparison of results provided interesting insight into how morphological properties of inhomogeneous particles determine their radiometric properties. The main findings were as follows:

- An external mixture model (Fig. 3c), in which LAC and sulphate are modelled as physically separated homogeneous spheres, can underestimate the absorption cross section by more than 100 %. This can be understood as follows. By adding a sulphate coating, the geometric cross section of a bare aggregate is increased, thus increasing the rate of electromagnetic energy that intercepts the particle and that is focused onto the LAC core, where part of the energy becomes absorbed. This effect is neglected in an external mixture model.
- A concentric core-shell model (Fig. 3e) with an LAC core and a sulphate shell underestimates absorption by up to 100 %, but only for the largest size parameters considered. For any given size parameter, the external mixture model underestimates the absorption cross section more strongly than the core-shell model. To understand this result, consider an electromagnetic wave interacting with an LAC sphere. If the size parameter of the sphere is sufficiently large, then, owing to the strong absorption, the electromagnetic wave will not fully penetrate to the centre of the sphere. Thus part of the LAC mass residing inside the sphere will not contribute to the absorption cross section. By contrast, in an LAC aggregate the carbon mass is distributed in such a way that a larger fraction of the mass comes in contact with the incident wave and absorbs electromagnetic energy. This explains why the encapsulated aggregate model predicts a

larger absorption cross section than the core-shell model and the externally mixed sphere model. Since the core-shell model does take the focusing effect into account, the underprediction of absorption is less pronounced in that model than in the external mixture model.

- A homogeneous "grey sphere" model (Fig. 3d) with an effective refractive index determined from EMA can overpredict the absorption cross section by almost 100 %. This model assumes sulphate and LAC to form a homogeneous mixture. As a result, essentially all of the LAC mass is in contact with the external field and contributes to absorption. By contrast, in an aggregate at least *some* of the LAC mass is shielded from interacting with the external field, although not as much as in a compact homogeneous sphere.

These results suggest that the most important morphological characteristic is the way in which the LAC mass is distributed inside the sulphate host particle. More precisely, it is important that a model particle correctly mimics the amount of LAC mass that interacts with the incident field, which determines the energy-absorption rate. This morphological feature makes the dominant contribution to the absorption properties of encapsulated LAC aggregates.

Having attained a more profound comprehension of the physics, one can devise improved model particles that incorporate our understanding of the relation between particle geometry and the light scattering process in these types of particles. Most importantly, a simple yet accurate model particle should yield an absorption cross section that lies between the extreme cases of the core-shell model (which underestimates absorption) and the homogeneous grey sphere model (which overestimates absorption). Thus in Kahnert et al. (2013) a "core-grey-shell" particle has been proposed (Fig. 3f). This particle is based on the simple core-shell geometry; however, only part of the LAC mass is concentrated in the core, while the remaining mass is homogeneously mixed with the sulphate shell. The "grey shell" composed of an LAC–sulphate mixture has an effective refractive index that is modelled with an EMA. Compared to the conventional core-shell model, the smaller core results in a smaller amount of the LAC mass being shielded from the external field, thus giving a larger absorption cross section. Compared to the homogeneous grey sphere model, a larger amount of the LAC mass is shielded from the external field, thus giving a lower absorption cross section. The amount of LAC mass that is concentrated in the core is a free parameter in this model. Note that the homogeneous grey sphere model and the conventional core-shell model are included in this model as limiting cases. It was found that this model reproduces not just the absorption cross section but also the single-scattering albedo and the asymmetry parameter over the entire range of particle sizes and wavelengths considered in that study (Kahnert et al. 2013). Even differential and spectral properties relevant to remote sensing, such as the backscattering cross section and the absorption Ångström exponent, were substantially improved in comparison to the other simple model particles. A qualitative summary of the performance of these models for different integral optical properties is given in Table 1. More quantitative details can be found in Kahnert et al. (2013).

Table 1 Qualitative performance of different homogeneous and inhomogeneous, but spherically symmetric model particles in reproducing the absorption cross section (C_{abs}), single-scattering albedo (SSA) and asymmetry parameter (g) of encapsulated LAC aggregates

Model	C_{abs}	SSA	g
Homogeneous sphere/external mixture	−−	++	−
Homogeneous sphere/internal mixture	+	−	+
Core-shell	−	+	−
Core-grey-shell	0	0	0

Plus and minus signs signify an over- and underestimation, respectively, of the reference results. Doubled plus or minus signs indicate a strong deviation over an extensive range of sizes and wavelengths. Zero indicates a generally good agreement

This example illustrates how the use of numerical laboratories can provide us with a more thorough understanding of the light-scattering process in particles with complex geometries, and how we can make use of such knowledge in developing simplified model particles that perform better than those that are merely based on ad hoc assumptions.

On the other hand, simple, spherically symmetric model particles rarely provide a faithful representation of the Mueller matrix of encapsulated LAC aggregates. As an example, Fig. 5 displays the phase function P_{11} and the degree of linear polarisation $-P_{12}/P_{11}$ for the model particles we just discussed. The particle size is 2 μm and the wavelength is 533.2 nm. The reference model of encapsulated aggregates is represented by the thick solid line. All models yield oscillations that agree qualitatively well with the reference case. This is because the number of oscillations is mainly dependent on the size parameter. The bottom row shows that the core-grey shell model (blue) gives the best overall representation of the phase function. Also, it reproduces the negative branch of the linear polarisation near the backscattering direction reasonably well. However, over the whole range of scattering angles, none of the models provides a good fit of the exact position and amplitude of the peaks and troughs of the linear polarisation.

Another recent study has considered encapsulated aggregates and bare aggregates to assess the impact of coating, volume fraction and fractal dimension on coated and bare aggregates (Cheng et al. 2014). The authors also considered a model of aggregates composed of thinly coated LAC spherules to represent LAC aerosols with a thin coating of weakly absorbing material. The results showed that the coating can enhance the absorption cross section of bare aggregates by up to a factor of 2, and the single-scattering albedo by almost a factor of 4, where the enhancement increases with the volume fraction of the weakly absorbing coating.

Apart from encapsulated geometries, model particles in which the LAC particle is attached to the surface of a weakly absorbing aerosol have been investigated. For instance, in Fuller et al. (1999), Mishchenko et al. (2004) a simple model consisting of an LAC sphere attached to a sphere of weakly or nonabsorbing material has been considered. More elaborate models consisting of a fractal LAC aggregate attached to a spherical sulphate or dust particle have been studied in Liu and Mishchenko (2007). Such geometries would result from coagulation processes. Note, however, that LAC

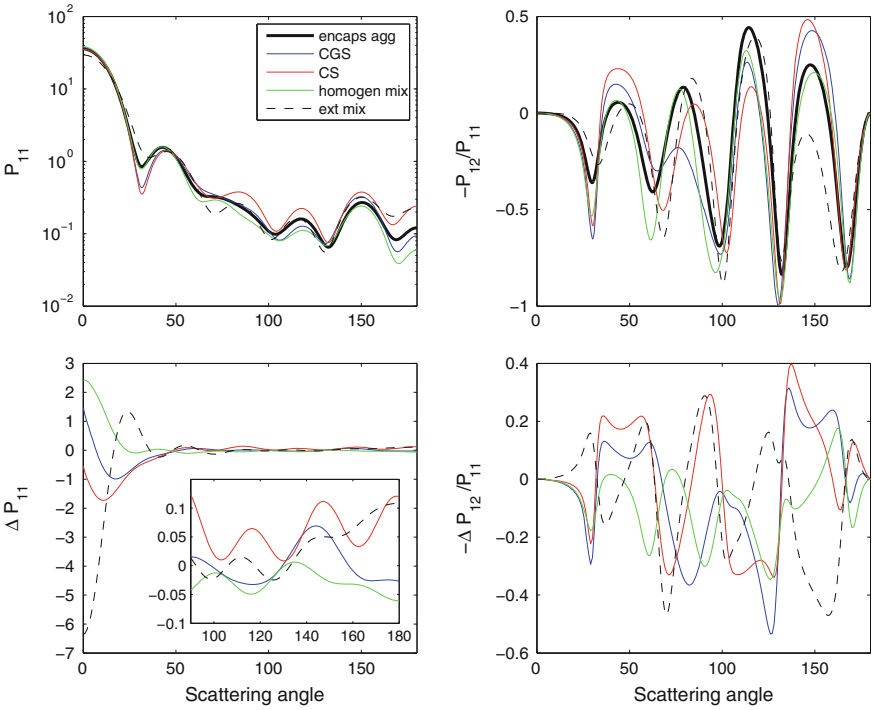


Fig. 5 Phase function P_{11} (*top left*) and degree of linear polarisation $-P_{12}/P_{11}$ (*top right*) computed for the encapsulated aggregate model (*solid black*), core-grey-shell (*blue*), core-shell (*red*), homogeneous grey sphere/Maxwell-Garnett (*green*), and external mixture model (*dashed black*). The difference between the simplified models and the encapsulated aggregate model are shown in the *bottom row*

and sulphate particles are mostly found in the Aitken and accumulation mode, while dust particles are mostly found in the accumulation and coarse mode. For particles in those size ranges, the coagulation coefficient is quite small; condensation is the dominant internal mixing process. Thus coagulated geometries with LAC particles attached to the surface of other aerosol particles are not very common; most aged LAC particles are partially or completely encapsulated by a shell of liquid material.

The aggregate models discussed thus far all neglect certain morphological subtleties. For instance, in real aggregates the monomer radius a may vary over a size distribution (Wu et al. 2015b), the monomers may be nonspherical (Wu et al. 2015a), or the average monomer radius in an aggregate may increase with the total aggregate size N_s. However, as noted in Bond and Bergstrom (2006), a variation in a has little effect on the absorption cross section, as long as $a<40$ nm, which is usually the case for atmospheric LAC particles. Changes in a can, however, affect the single-scattering albedo (Kahnert 2010c). A potentially more severe approximation is the assumption that neighbouring monomers in the aggregate are touching in a single point only. In reality, the geometry is more likely to be sintered;

i.e., there are bridges between neighbouring spheres. It is known from model calculations on titanium dioxide aggregates that sintering can significantly enhance the extinction cross section, especially at UV wavelengths (Wriedt et al. 2010). Modelling studies on sintered and non-sintered soot aggregates have also been reported (Skorupski et al. 2013), but not for soot aggregates that would be representative for atmospheric LAC.

Finally, we shall briefly discuss computational methods used for modelling light scattering by LAC particles. Some modelling studies of LAC aggregates are based on non-rigorous solutions to Maxwell's equations, such as the Rayleigh-Debye-Gans (RDG) approximation (Bohren and Huffman 1983; Zhao and Ma 2009). This approach neglects electromagnetic interaction among the monomers in the aggregate. In Xu and Gustafson (2001), a generalisation of Mie theory to aggregates of multiple spheres was developed. The method can be used as a numerically exact solution to Maxwell's equations, but it also allows for approximate solutions. To this end, an interaction index $f = (a_i + a_j)/|r_i - r_j|$ is introduced, where a_i and a_j are the radii of monomers i and j and $|r_i - r_j|$ represents their separation. Electromagnetic interaction among monomers with a value of f lower than a prescribed value is neglected. A prescribed value of $f = 0$ corresponds to the exact solution, while $f = 1$ completely neglects interaction among all monomers. This method has been applied to LAC aggregates in Kahnert (2010c), where it was shown that neglecting electromagnetic interaction among the monomers can result in large errors in modelled absorption cross sections. Thus it is not generally guaranteed that RDG-type approximations are sufficiently reliable for modelling light scattering in LAC aggregates.

The most commonly used numerically exact methods for performing light-scattering computations on LAC aggregates are the superposition T-matrix method (STMM) (Mackowski and Mishchenko 1996) and the discrete dipole approximation (DDA) (Purcell and Pennypacker 1973). The accuracy of the DDA has been confirmed by performing comparisons of results for bare aggregates computed with the DDA and the STMM (Kahnert et al. 2012), and by testing the reciprocity condition (Schmidt et al. 2012) for bare and encapsulated aggregates (Kahnert et al. 2012). There exist well-tested open-source implementations of the DDA, such as DDSCAT (Draine and Flatau 1994; Draine 2000) and ADDA (Yurkin and Hoekstra 2011). The DDA can be applied to arbitrary shapes. Until recently, the STMM could only be applied to configurations of spheres that are external to other spheres. Thus applications of the method to LAC particles were limited to bare aggregates. Recently, a new implementation of the multiple-sphere T-matrix (MSTM) code has been released (Mackowski and Mishchenko 2011), which has been generalised to arbitrary configurations of spheres located internally or externally to other spheres, with the only restriction that the surfaces of the spheres must not overlap. This allows us to apply the MSTM code to LAC aggregates internally mixed with weakly absorbing aerosol components. Note that in encapsulated geometries part of the aggregate can be sticking out of the liquid-phase host particle. In the DDA method, it is possible to treat geometries in

which individual monomers only partially stick out of the surface of the host, such as in the model particle shown in Fig. 3b. The non-overlap restriction in MSTM requires that each LAC monomer is either completely inside or completely outside the encapsulating host sphere, i.e. a monomer must touch the surface of any other sphere, including the host, in not more than a single point. However, it has been shown that the extinction cross section, single-scattering albedo and asymmetry parameter of semi-embedded LAC-containing mixtures with non-overlapped and overlapped morphologies differ by less than 1 % (Wu et al. 2015c).

4 Mineral Dust

Mineral dust is one of the most prominent aerosol types in the Earth's atmosphere. It consists of soil-constituting mineral particles originating mainly from arid and semi-arid regions through weathering and erosion processes. Deserts, in particular, are important sources of mineral dust. Dust particles exhibit a wide variety of shapes and may be inhomogeneous aggregates or agglomerates of different minerals; just as snow flakes, each dust particle is unique.

There are a wide variety of ways to attain information about the shape and structure of dust grains. For example, scanning (SEM) and transmission (TEM) electron microscopes can be used to image dust particles in the nanometre scale. Electron microscopy also yields data on the spatial distribution of particle composition, such as with the energy-dispersive X-ray spectroscopy (EDS) mapping in SEM and TEM. The latter provides data also on the atomic lattice structures, but requires the analysed target to be electron transparent, which super-micron dust particles generally are not. Recently, focused ion-beam (FIB) milling has been adapted for preparing cross-sectional slices from dust particles too thick to be electron transparent, to allow the analysis of the dust particle interior and its three-dimensional structure in more detail (Jeong and Nousiainen 2014). For dust particles with diameters in tens or hundreds of micrometres, optical microscopes can also be used. Figures 6 and 7 illustrate data attained on dust particle inhomogeneity with some of these techniques.

Thousands of different minerals have been identified from the Earth's soil. Most of them are, however, exceedingly rare, and the dust–aerosol mass is dominated by a few mineral families, namely quartz, phyllosilicates (mainly micas, clay minerals and chlorites, for example biotite, illite, kaolinite, montmorillonite or palygorskite), feldspars, carbonates (mainly calcite and dolomite), sulphates (gypsum/anhydrite, but also other evaporites), and oxide minerals (mainly goethite, hematite, rutile or ilmenite). On an individual particle level, these different components are anything but evenly distributed.

The dust particle shapes can be roughly divided into mono grains, main grain with minor adhesions, agglomerates, and aggregates (Nousiainen and Kandler 2015). Mono grains may show explicit crystal structure, but more often are rather irregular. Typically, their shapes are roundish and roughly equidimensional,

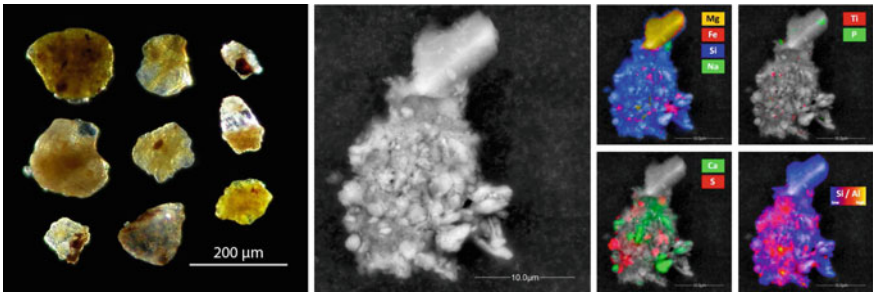


Fig. 6 Examples of dust particle inhomogeneity. The *leftmost panel* shows light microscope images of large dust grains, where iron(oxi)hydroxides are seen as *yellow* or *reddish* hue (Kandler et al. 2009). The *middle panel* shows a SEM image of a dust particle, whose compositional inhomogeneity is visualised in the four sub-panels at right (Scheuvens et al. 2011). The spatial distribution of different elements is visualised by the corresponding colour intensity. The Si/Al ratio in the *lower right subpanel* allows identifying changes in the silicate matrix (e.g., identifying quartz grains)

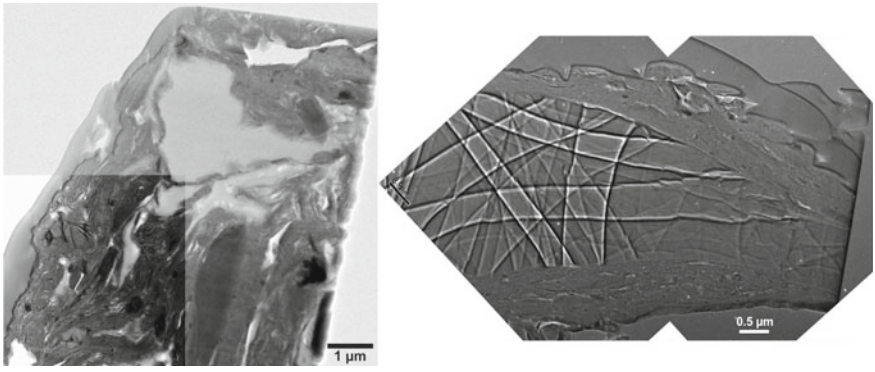


Fig. 7 Internal structures of two Asian dust grains observed with transmission electron microscopy from cross-sectional dust particle slices prepared with focused ion-beam milling (Jeong and Nousiainen 2014). The images show substantial degrees of inhomogeneities and internal structures, including pores

angular, or thin flakes or needles whose aspect ratios may even exceed 10, depending on their composition. Main grains with adhesions are mono grains on which some small, usually thin clay flakes have adhered. The agglomerates and aggregates are both formed from multiple grains without a dominating one. Their difference is in the strength of cohesion and compactness: agglomerates are looser and easier to break up due to mechanical stress.

As different minerals have varying hardness and breaking habits, particle sizes, shapes and compositions are linked: different minerals exhibit different size ranges and characteristic shapes. From an optical modelling point of view, this is an additional complication, because it means that, in principle, sizes, shapes and

compositions cannot be treated independently. So, for example, when solving ensemble-averaged optical properties of dust for remote-sensing or climate-modelling purposes, one should solve the optical properties separately for each combination of particle size, shape and composition. This is exceedingly laborious, and has never been done. Rather, the common practice is to assume a specific size distribution, choose a shape model and assign a bulk composition. These parameters are then applied throughout the particle ensemble. Errors thus introduced are, as of now, unassessed.

Modelling approaches for dust are many. In most climate models, they are still modelled as homogeneous spheres, with two recent exceptions (Wang et al. 2013; Räisänen et al. 2013). Other models for dust include, for example, spheroids (Mishchenko et al. 1995; Nousiainen and Vermeulen 2003; Dubovik et al. 2006), ellipsoids (Bi et al. 2009), polyhedral prisms (Nousiainen et al. 2006), non-symmetric hexahedra (Bi et al. 2010), convex polyhedra (Gasteiger et al. 2011), Gaussian random spheres (GRSs) (Nousiainen et al. 2003; Veihelmann et al. 2006), random blocks of cubes (Kalashnikova et al. 2005; Vilaplana et al. 2006), deformed spheroids and their aggregates (Gasteiger et al. 2011), concave fractal polyhedra (Liu et al. 2012), spatial Poisson–Voronoi tessellation (Ishimoto et al. 2010), agglomerated debris particles (Zubko et al. 2013), irregular flakes (Nousiainen et al. 2009), irregular rhombohedra (Dabrowska et al. 2012) and inhomogeneous stereogrammetric shapes (Lindqvist et al. 2014). In a stark contrast with the diverse set of model shapes considered, very little attention has been paid to the dust particle inhomogeneity. Indeed, almost all of the models used are homogeneous. Further, even when the inhomogeneity is accounted for, it is usually done either by using external mixtures (e.g., Gasteiger et al. 2011), or by applying an EMA (see Sect. 2) to replace the inhomogeneity by an effective homogeneous composition. Even when this approach is not explicitly applied, it is often implicitly used, as few studies actually use refractive indices of specific mineral species, but rather adapt a typical value for mineral dust at the wavelength of interest.

Almost all studies that explicitly account for the inhomogeneity of dust particles are sensitivity-type modelling studies that aim at investigating impacts of inhomogeneity on scattering by using dust-like parameter values, rather than attempting to model actual, observed inhomogeneity characteristics of real dust particles. For example, Vilaplana et al. (2006) studied the impact of porosity by randomly removing half of the volume elements describing the model particle. Mixtures of different minerals were not considered. Indeed, their motivation was rather in understanding astronomical observations of possibly quite fluffy cometary dust. They assumed the dust particles to be composed of rather absorbing mineral, and for such particles the introduction of internal porosity turned out to greatly alter the size dependence of scattering, in particular for the degree of linear polarisation and (linear) depolarisation.

In another similar study, Lindqvist et al. (2009) considered inhomogeneous mixtures of silicate-type materials and water ice. In their model, the inhomogeneity was modelled using a ballistic cluster of spheres, which was then wrapped in a coating modelled as a so-called concave hull. Additionally, they also considered silicate GRSs coated with ice. Their DDA simulations revealed similar, albeit

weaker effects to those obtained by Vilaplana et al., most likely because the dielectric contrast between silicate and ice is much weaker than between silicate and free space.

As a third example, Kahnert (2015) investigated the optical impact of the spatial distribution of hematite in a quartz particle. He used a spherical quartz grain with a varying volume fraction and number of spherical hematite inclusions within. For scattering computations, he used the STMM. These simulations revealed quite substantial impacts on scattering by as small as one vol% of hematite in the particles. The effect was larger for many small hematite inclusions than for one big inclusion. Interestingly, while Lindqvist et al. (2009) report good performance for EMAs in mimicking the impact of inhomogeneity in scattering computations, in Kahnert's study the opposite is true. This different performance may be attributable to the larger dielectric contrast between the components in the latter study, or to the fact that in the latter study one component was strongly absorbing and one was not, while in the former study neither component absorbed much. Finally, it is possible that some contribution comes from the use of spherical model particle in the latter study, while the former employed irregularly shaped particles.

The only investigations the authors are aware of that attempt to explicitly model the observed spatial distribution of mineralogical composition of real dust particles are those by Lindqvist et al. (2014) and Kemppinen et al. (2015). In the former study, stereogrammetric analysis of SEM images was used to derive the real three-dimensional shape for four individual dust grains, while the SEM-EDS analysis was used to map the distribution of component minerals in the particles' exterior. As the interiors of the particles could not be analysed explicitly, assumptions were used to extend the mineralogical features derived by the observed surface morphological features and corresponding elemental compositions into the particles' interiors. In the latter study, the three-dimensional Voronoi tessellation algorithm was applied to generate irregular, inhomogeneous model particles. The parameters for the model were chosen to generate similar internal structures as seen in real dust particle interiors (Jeong and Nousiainen 2014). Thus, Kemppinen et al. used a mathematical model to mimic the real, observed internal structure and inhomogeneity of dust particles and give the particles some kind of irregular overall shape; whereas, Lindqvist et al. derived directly the shape and surface composition of real dust particles, but had to assume their internal characteristics. In both studies, the single-scattering properties were solved using the DDA for a range of size parameters without making any simplifications to the derived microphysical characteristics of the target particles. The model shapes used in these studies are illustrated in Fig. 8.

Lindqvist et al. found that, for their model particles, the inhomogeneity did not induce substantial effects on scattering when the dielectric properties of the component minerals were similar. However, even two vol% of strongly absorbing hematite introduced considerable effects. They also found that the tested EMAs did not perform well in accounting for the impact on scattering and absorption from the hematite. Likewise, Kemppinen et al. found that EMAs generally failed to account for the impact of inhomogeneity on scattering, except when all components had

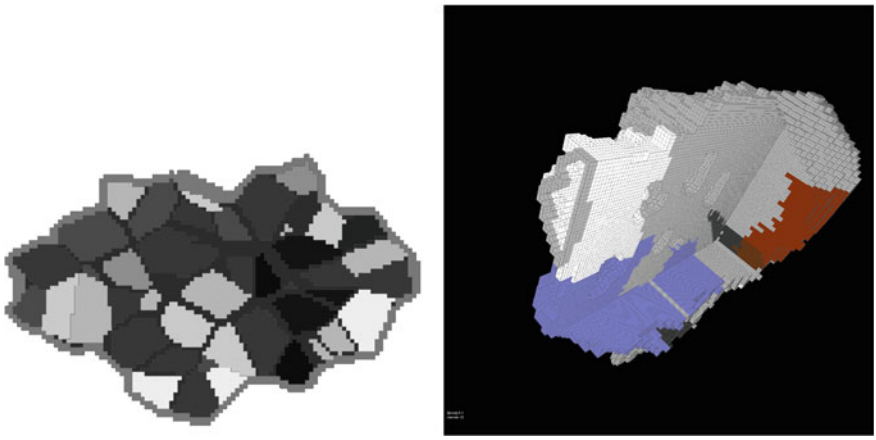


Fig. 8 Examples of inhomogeneous model dust particles. A cross section of a three-dimensional Voronoi-tessellation particle from Kemppinen et al. (2015) (*left*), and stereogrammetrically derived 3D shape with SEM-EDS analysis-based inhomogeneity from Lindqvist et al. (2014) (*right*). Note that the latter has a segment removed to reveal the internal structure

very similar refractive indices. They also found that different types of inhomogeneity studied, e.g. internal pores, hematite nodes and coatings, had clear but different impacts on scattering. To account for their impacts on scattering accurately, it appears necessary to explicitly model them.

Some effects of inhomogeneity on scattering by model dust particles are illustrated in Fig. 9. The top panel illustrates the effect of hematite inclusions on three scattering matrix elements of a spherical, nonabsorbing silicate particle with size parameter $x = 8.1$. The black line presents a case where the impact of 100 hematite inclusions with the total volume of 4 vol% of the particle has been explicitly accounted for in the simulations. The green line shows the corresponding case, where the impact of the inclusions have been accounted for by using an effective refractive index derived by computing the linear average of volume-weighted refractive indices of the components. In this case, the explicit accounting for the inhomogeneity has a drastic effect on the P_{22}/P_{11} scattering matrix element; the inclusions give rise to anisotropy, seen as a deviation from unity for this quantity, which cannot be mimicked by any kind of isotropic, homogeneous composition. Other scattering matrix elements also show considerable differences, with the inhomogeneous case showing generally much weaker interference structures in the angular dependence of scattering. The lower panel shows results from a similar investigation with an irregular and inhomogeneous base shape. The baseline case, shown with blue dashed line, shows the angular dependence of the scattering matrix elements for an ensemble of particles such as that shown in the left panel of Fig. 8, with size parameter $x = 8$ and mineral composition consisting of different clay minerals with similar, nonabsorbing refractive indices. The black line presents simulation results for corresponding particles with 17 vol% of spherical hematite

inclusions randomly located within the particles, while the dashed red line presents a case where both all the clay minerals and the added hematite inclusions have been averaged into a homogeneous, effective composition using, again, the linear average of volume-weighted refractive indices of the components. The figure shows that the impact of hematite inclusions on scattering cannot be mimicked accurately with an effective, homogeneous medium in this case either, but the differences are not as drastic as for the spherical particles in the top row, even though these particles contain much more hematite.

It thus appears that the impact of inhomogeneity on scattering depends on the type of host particle, as illustrated in Fig. 9: inhomogeneity adds to the complexity of the target particle, and is therefore likely to have more substantial impact on scattering when occurring in otherwise simple model particles such as spheres. That is to say, the single-scattering properties are likely to differ more between homogeneous and inhomogeneous spherical particles, than between homogeneous and

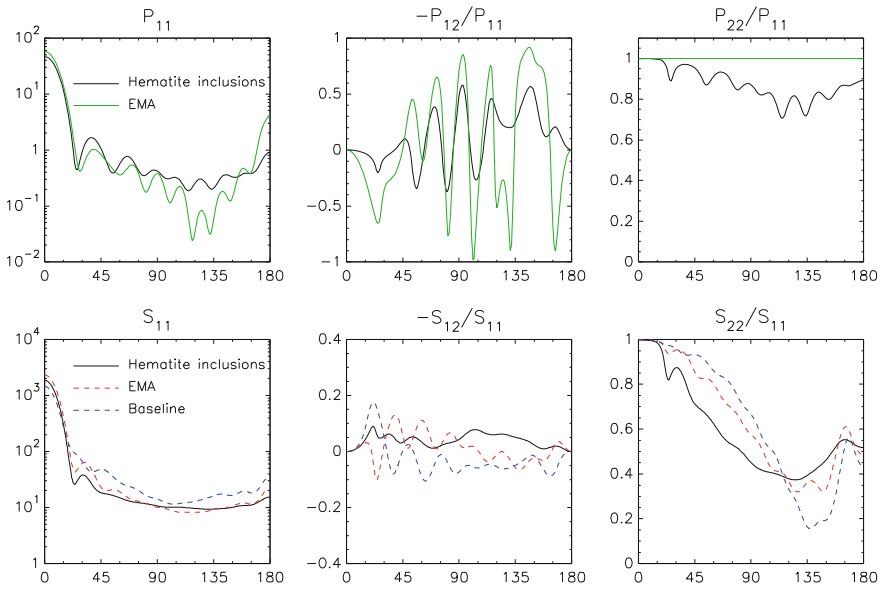


Fig. 9 Illustration of the impact of inhomogeneity on the angular dependence of scattering by model dust particles. The *top row* presents results for a spherical silicate particle with added spherical hematite inclusions with data taken from Kahnert (2015). The *black line* presents the inhomogeneous case and the *green line* a homogeneous case with the effective refractive index derived with an EMA. The *bottom row* shows the corresponding results for irregular, inhomogeneous particles generated with three-dimensional Voronoi tessellation (e.g., Fig. 8, *left panel*), with the data taken from Kemppinen et al. (2015). The *dashed blue line* shows a baseline case composed only of nonabsorbing clay minerals. The *black line* shows the results after the particles have been ingrained with spherical hematite inclusions, while the *dashed red line* shows the corresponding homogeneous case with the effective refractive index derived with an EMA. Note that the top and bottom rows use differently normalised scattering matrices, so the *y* axis values in the leftmost panels are not comparable

inhomogeneous complex-shaped particles. In general, inhomogeneity tends to give rise to scattering characteristics such as flat phase function and weak, positive linear polarisation at side-scattering angles, which are also associated with scattering by complex, homogeneous particles, while spherical particles generally scatter light considerably differently. Naturally, in each case the effects will weaken as the dielectric contrast diminishes.

5 Volcanic Ash

Volcanic ash originates mainly from volcanic eruptions where magma cools rapidly either by explosive decompression (magmatic/phreatic eruption) or due to contact with water or ice (phreatomagmatic eruptions). Most ash particles are formed through bubble fragmentation, when magma rises rapidly to the surface, and the decreasing pressure releases gases dissolved in the magma (Durant et al. 2010). In the process, bubbles or vesicles are often formed within the ash particles, making them porous. Micrometre scale and smaller ash particles are often shards from the bubble walls. Similarly to mineral dust, ash particles present widely varying shapes. In the absence of erosive morphing, however, they often exhibit sharp edges. Example ash particles are shown in Fig. 10. The high level of porosity of these particles is particularly clear.

Besides porosity, ash particles can be inhomogeneous in mineral composition. As the particles in a single ash cloud originate from the same magma source, ash clouds are typically less heterogeneous than desert dust plumes (Nousiainen and Kandler 2015; Schumann et al. 2011). This does not necessarily mean, however, that single ash particles would be less inhomogeneous than individual dust grains. For example, analyses of ash particles from the 2011 Grímsvötn eruption by Lieke et al. (2013) show individual ash particles composed of multiple minerals; distinct enrichments of Ti, Fe and Mg were found in different parts of the particles. In particular, crystals with high melting point, such as iron and titanium oxides, may

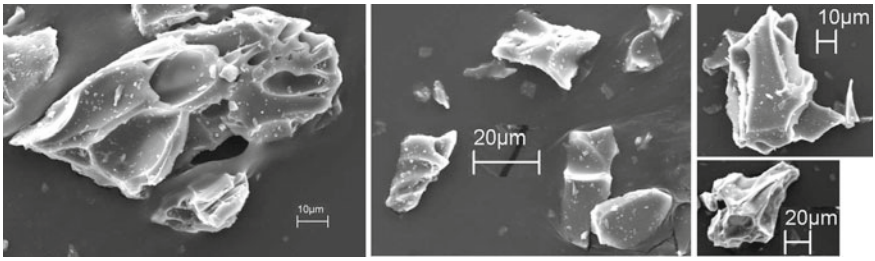


Fig. 10 Example SEM images of ash particles from the 2011 eruption of the Puyehue-Cordón-Caulle complex, Argentina

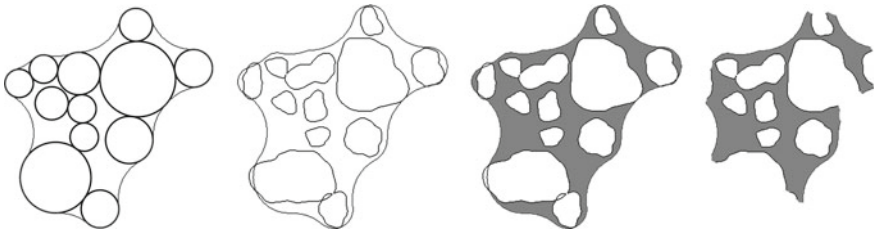


Fig. 11 Two-dimensional cross sections illustrating the steps in the generation of model ash particles by Lindqvist et al. (2011)

remain in crystalline form in the melt and thus form inhomogeneities in the resolidified particles.

Volcanic ash particles have been modelled, e.g. as homogeneous spheres (Niemeier et al. 2009), spheroids (Krotkov et al. 1999), ellipsoids (Merikallio et al. 2015) and nonsymmetric hexahedra (Bi et al. 2010). Again, few models have been developed to explicitly account for the inhomogeneity, and even they consider only inhomogeneity due to porosity. The most detailed, albeit still descriptive rather than directly retrieved shape model is that by Lindqvist et al. (2011), where the ash particle porosity and irregular shape are accounted for. Visual illustration of the generation process, for a two-dimensional cross section, is shown in Fig. 11. Generation of their model particles begins by creating a dense ballistic cluster of spheres with a given total number and size distribution. Then, the cluster is enveloped in a concave surface by the concave-hull transformation (Lindqvist et al. 2009), where the radius of the so-called generating sphere presents another free parameter for the shape model (leftmost panel). Then, each sphere in the cluster is replaced by non-spherical GRSs with user-defined shape parameters σ and v, which has the same origin as the spheres in the cluster, to describe non-spherical porous cavities or vesicles in the particle (second panel). The volume between the concave hull and the vesicles is then filled (third panel). Finally, a surface layer of a given depth is peeled from the particle to expose some of the cavities, creating a cratered surface (rightmost panel). The generation is a random process, so different model ash particles can be generated with each set of parameters simply by changing the seed for the random number generator. The single-scattering simulations carried out with the DDA at 633 nm wavelength showed that the porosity of ash particles promote positive linear polarisation at side-scattering angles, which is consistent with laboratory-measured scattering by ash particles (e.g., Muñoz et al. 2004). The impact was larger with many small vesicles in the particle, compared to fewer larger vesicles. Both types of vesicles tended to reduce the strong depolarisation caused by the host particle without vesicles, but for small vesicles the effect was much stronger. Example results shown in the lower panel of Fig. 12 illustrate these general effects, except that for the chosen size parameter the large-vesicle case did not show markedly different depolarisation (P_{22}/P_{11}) to that by the solid ash particle.

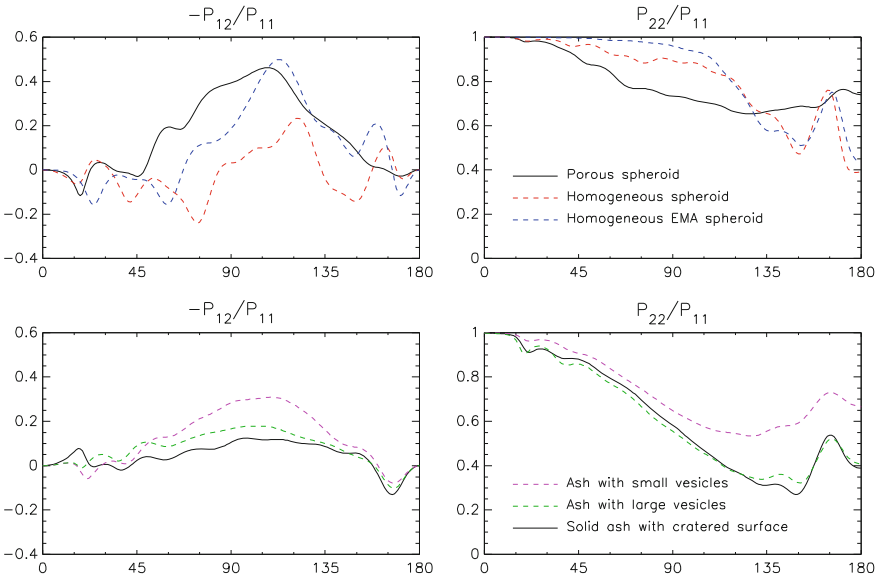


Fig. 12 Impact of inhomogeneity on scattering for modelled ash particles at size parameter $x = 8$. The *top row* shows scattering matrix elements $-P_{12}/P_{11}$ and P_{22}/P_{11} for oblate spheroids. The *black solid line* corresponds to an aspect-ratio 2.0 spheroid with small cavities (Fig. 12, *right panel*). The *dashed lines* show results for the corresponding solid spheroids; *red* with the matching refractive index, and *blue* with a Maxwell Garnett effective refractive index to account for the cavities. Data are taken from Nousiainen et al. (2011). The *bottom row* shows results for irregular vesicular ash particles by Lindqvist et al. (2011) (e.g., Fig. 11). The *magenta dashed line* corresponds to an ash particle with small vesicles (porous cavities), the *green* one with large vesicles, and the *black solid line* a case where the internal vesicles have been filled, but the exterior of the particle remains cratered

Another porous model is that of porous spheroids (Fig. 13) introduced in Nousiainen et al. (2011). For that model, too, the porous cavities are created using a ballistic cluster of spheres, only in this case a spheroidal shape is cut out of the porous medium, and the spherical cavities are not replaced by non-spherical shapes. Even though this shape model was not generated with volcanic ash in mind originally, it was later used to model ash infrared single-scattering properties by Kylling et al. (2014), together with the shapes by Lindqvist et al. (2011). Kylling et al. (2014) found that, when compared against spherical ash particles, the use of non-spherical model particles resulted in detectable ash signature in the brightness temperature (different brightness temperature at 10.8 and 12 μm wavelengths) for larger ash particles. The use of spherical, homogeneous ash particles in satellite retrievals might thus explain why satellite measurements never show as large ash particle sizes as those found in distant ash sediments (Stevenson et al. 2015). In addition, spherical model particles were found to underestimate the mass loading, compared to the use of irregular model particles, in thermal infrared retrievals. The effects of porosity on light scattering by these spheroids at the visible wavelength of

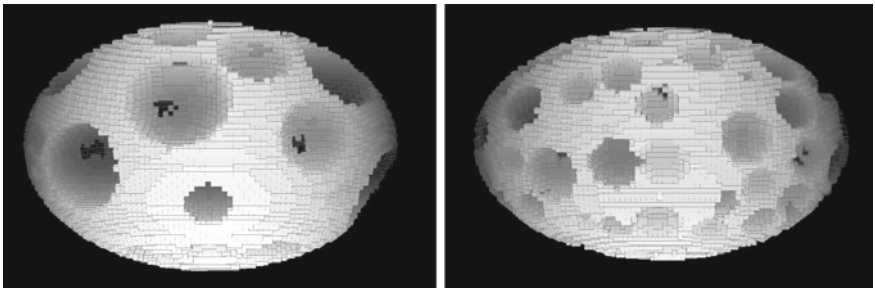


Fig. 13 Example images of porous spheroids with large (*left*) and small (*right*) cavities developed by Nousiainen et al. (2011) and used for modelling ash particle infrared signatures by Kylling et al. (2014)

$\lambda = 628$ nm are illustrated in the top row of Fig. 13. It can be seen from the plots that the presence of internal pores dampens the oscillations caused by the symmetric, spheroidal base shape. The pores also tend to make the $-P_{12}/P_{11}$ more positive. Finally, it is obvious that the EMA is unable to account for the impact of porosity of the particles.

6 Cometary Dust

Cometary dust is believed to be an unequilibrated heterogeneous mixture of various components which possibly do not share a common origin. Dust particles may contain presolar particulates as well as solar nebula condensates formed in the hot inner and the cold outer regions of the protosolar disk (Hanner and Bradley 2004; Zolensky et al. 2006). Subsequently, these constituents may have aggregated to form larger highly porous particles and incorporated into comets.

Infrared spectroscopy, in situ measurements and analyses of interplanetary dust particles (IDPs) suggest that cometary dust particles are porous heterogeneous mixtures of rock-forming elements such as Mg, Si, Ca, Fe and carbonaceous materials so-called CHON-component consisting of light elements: carbon, hydrogen, oxygen, and nitrogen lawler. These two major constituents are thought to be mixed at sub-micrometre scale, possibly containing micrometre-sized mineral grains coated or attached together by carbonaceous material. The rock-forming elements and the CHON-materials have been found by in situ sampling of comets 1P/Halley and 81P/Wild 2 carried out by the two Vega spacecrafts, ESA's Giotto probe and NASA's Stardust sample return mission (Zolensky et al. 2006; Lawler and Brownlee 1992; Jessberger et al. 1988; Brownlee et al. 2006; Sandford et al. 2006). In addition, these materials are known to be the major constituents of IDPs captured in the Earth's stratosphere; consequently, IDPs are suggested to be of cometary origin (Bradley 2004; Ishii et al. 2008). Laboratory measurements of IDPs

have exposed a composite structure with inhomogeneity on the sub-micrometre scale (Bradley 2004; Bradley et al. 1992; Keller et al. 2000). These findings are also consistent with the model of cometary dust formation by Greenberg (1982), Greenberg and Hage (1990). The model suggests that the dust consists of aggregated core-mantle particles with a silicate core, an organic refractory mantle and possibly an outer coating of ice. Despite the above-mentioned in situ observations, there are no direct measurements of the internal structure of cometary dust available at the moment. The Stardust mission samples were affected by a rapid deceleration in the capturing process, and thus, the internal structure of particles were mostly destroyed. The Rosetta spacecraft by ESA has successfully rendezvoused with the comet 67P/Churyumov-Gerasimenko in August 2014 and is now in orbit performing detailed analysis of the dust environment of the coma.

Despite the heterogeneity of cometary dust particles, they are often modelled as homogeneous shapes with a constant refractive index (see e.g. Kolokolova et al. 2004; Mukai et al. 1991; Xing and Hanner 1997; Petrova et al. 2000; Kimura et al. 2006 and references therein). The refractive index is usually obtained by some mixing formula, e.g., Maxwell Garnett or Bruggeman or left as a free parameter. Particle shapes are most often constructed by packing identical spheres using the ballistic cluster-cluster BCCA or ballistic particle-cluster BPCA aggregation processes. However, the dust grains can vary in size and probably are not spherical. Hence, models containing aggregated spheroids or GRSs have been used to decrease morphological regularity (Lumme and Penttilä 2011). In addition, highly irregular shapes, so-called "agglomerated debris particles" have been applied (e.g. Zubko et al. 2011). The use of a constant refractive index, however, may not be sufficient at the visible wavelengths since the microscopic images of IDPs suggest that size of the constituent grains are comparable to the visible wavelengths. Moreover, the refractive indices of silicate-type minerals and carbonaceous materials differ significantly. Silicates are almost transparent whereas CHON-materials are highly absorbing at visible wavelengths, indicating that the light-scattering properties can be sensitive to the internal composition of the particles. This suggests that inhomogeneities should be taken into account in the analysis of light scattering by cometary dust.

Only a few models explicitly account for inhomogeneous composition of the particles. In Yanamandra-Fisher and Hanner (1999), light scattering by small heterogeneous (two-component) aggregates were studied with the DDA. Inhomogeneity was introduced by randomly choosing dipoles in the DDA grid and changing the refractive index of the particular cell, thus leading to a small-scale mixture since the DDA cells are always small compared to the wavelength. It was concluded that the small-scale mixture of silicate and carbon is not suitable for cometary dust models since even a small amount of carbon embedded in the silicate matrix weakens the negative polarisation and the enhancement of the intensity near the backscattering direction, which are typical light-scattering characteristics observed for cometary dust. Figure 14 shows the degree of linear polarisation in the narrow-band blue filtre versus phase-angle for various comets. Larger spheroidal graphite inclusions in silicate host spheroids were studied in Gupta et al. (2006); however, the authors could not reproduce the negative polarisation at small phase

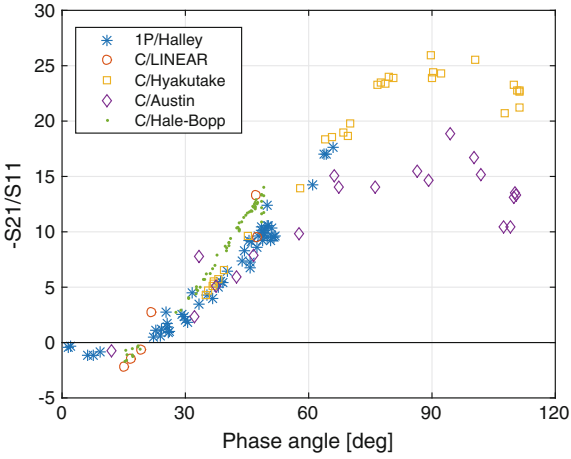


Fig. 14 Polarimetric observations for different comets at $\lambda = 484.5$ nm (Kiselev et al. 2006). The observations show a shallow branch of negative polarisation near the backscattering region and a positive branch with a maximum near 90–100°

angles either. In Greenberg's grain-formation model (Greenberg 1982; Greenberg and Hage 1990), the dust grains are submicrometre-sized silicate cores covered by organic refractory mantles. These grains were used as compact constituents of porous BCCA or BPCA aggregates in Lasue and Levasseur-Regourd (2006), Levasseur-Regourd et al. (2007), but the scattering properties of these grains were found to be very similar to those of the pure organic grains assuming that the organic mantles are sufficiently thick. In combination with larger compact spherical grains, the authors were able to reproduce the polarimetric observations of the comets 1P/Halley and C/1995 O1 Hale-Bopp for the positive-polarisation part, but they did not obtain a good fit for the negative polarisation branch (Lasue et al. 2009).

Mixtures of agglomerated debris particles were used to fit the polarimetric data of comet C/1975 V1 West in Zubko et al. (2014). However, only external mixtures of pure silicate and pure amorphous carbon particles were considered. Inhomogeneous irregular particles with the agglomerated structure were introduced in Markkanen et al. (2015) and their scattering properties were compared with external mixtures of homogeneous particles as well as particles whose refractive indices were obtained by the Maxwell Garnett mixing formula. Figure 15 demonstrates the difference between the scattering properties of the inhomogeneous model and the Maxwell Garnett EMA. The generation process of the inhomogeneous irregular particle is illustrated in Fig. 16. First, a sphere is meshed with e.g. tetrahedral/cubical elements (depending on the numerical solver). Then seed points are randomly chosen inside the sphere, and the corresponding Voronoi cells are found, i.e., each Voronoi cell contains all tetrahedral elements whose centres are closer to that seed point than any other seed. Each Voronoi cell can be associated with a chosen refractive index or a subdomain. If the refractive index of a particular Voronoi cell is the same as in background, it can be removed. Moreover, the resulting particle can be used as a starting point for the finer level Voronoi partitioning allowing hierarchical control of inhomogeneity, porosity and irregularity in a statistical sense.

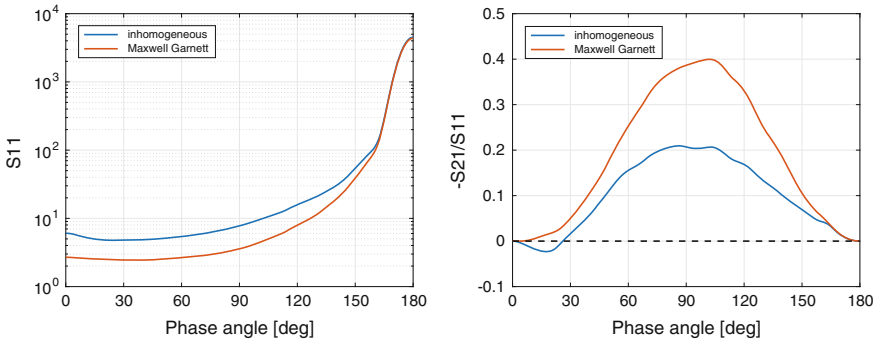


Fig. 15 Comparison between the inhomogeneous model and the Maxwell Garnett EMA. The phase function and the degree of linear polarisation are computed for an ensample of irregular particles. The average volume-equivalent size parameter of the particles is 8 and the average size of grains constituting the particles is 1.6. The particles consist of 30 % amorphous carbon ($m = 2.45 + 0.45i$) and 70 % silicates ($m = 1.7 + 0.003i$) in volume

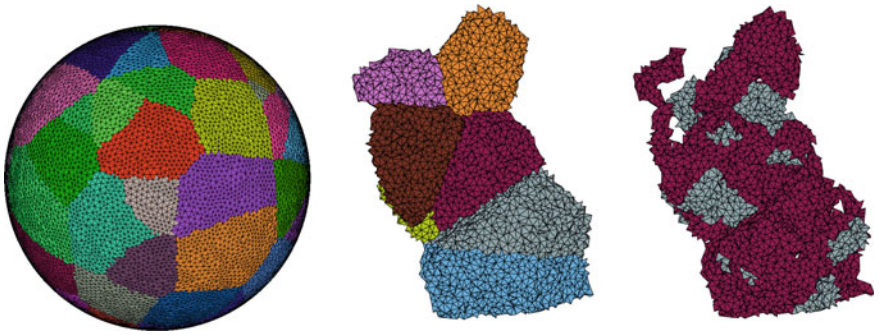


Fig. 16 Illustration of the generation of level-two inhomogeneous Voronoi particle on tetrahedral mesh. A sphere is meshed with tetrahedral elements and the mesh is partitioned into Voronoi cells (*left*). Randomly chosen cells are removed and the level-one irregular shape is obtained (*middle*). Level-two Voronoi partitioning is executed and material parameters (two different materials) are randomly chosen for each level-two Voronoi cell (*right*)

7 Biological Particles

In biomedical optics, the incentive for considering inhomogeneous morphologies is rather different from that in atmospheric optics. In the atmosphere the physical and chemical processes of interest influence mainly the size and chemical composition of aerosols. On the other hand, morphology is not a property that is accounted for in aerosol dynamic and atmospheric chemical transport models (apart from a simple description of the mixing state of different chemical compounds). Therefore, to account for morphological complexity in aerosol optics models is often an

inconvenience that one tries to avoid whenever possible. (This may, however, change as soon as retrieval methods of quantities, such as lidar backscattering depolarisation, become sufficiently mature to obtain information on aerosol composition from the aerosol morphology.) By contrast, in biomedical optics tissue morphology is often the main property of interest that one wants to derive from light-scattering measurements. The aim is to develop non-invasive methods for evaluation of tissue pathology by using experimental techniques, such as diffuse optical spectroscopy, elastic scattering spectroscopy, or diffraction imaging flow cytometry. Since carcinogenic cells differ from healthy cells mainly in the internal structure, one needs to attain a thorough understanding of the relationship between cell morphology and light-scattering characteristics.

The cell wall contains the cytoplasm and various cell organelles. The cytoplasm accounts for some 40–80 % of the cell volume, the nucleus for 5–10 %, the mitochondria for 5–15 %, and other organelles for 1–10 % (Alberts et al. 1989). Typical cells have sizes in the range 10–30 μm, but they can be as small as some nanometres. The nucleus ranges between 3 and 10 μm and is, therefore, the largest organelle. Mitochondria range between 0.5 and 1.5 μm, endoplasmatic reticulum between 0.2 and 1 μm, lysomes between 0.2 and 0.5 μm, and peroxisomes between 0.2 and 0.5 μm. Refractive indices for various cell constituents have been compiled from different literature sources by Dunn (2007). In comparison to aerosols or cosmic dust particles, biological particles are optically rather soft. Both the dielectric contrast between the cytoplasm and empty space, and the dielectric contrast between the cell organelles and the cytoplasm are quite small. Under certain circumstances this may allow for the use of approximate theories, such as the RDG approximation. However, it has been shown in Orlova et al. (2008) that RDG theory as well as the second order Born approximation have their limitations even when applied to biological particles. For this reason, several authors have relied on the use of numerically exact methods for solving the electromagnetic scattering problem, such as the DDA (Orlova et al. 2008) or the finite-difference time-domain method (Dunn 2007).

The usefulness of drastically simplified model geometries, such as homogeneous and coated spheres, is fairly limited in biomedical modelling (Starosta and Dunn 2010). To understand the impact of cell morphology on optical properties one needs to account for the inhomogeneous cell structure. However, the cell membrane seems to have only a small effect on scattering properties, so it can usually be ignored in the construction of model particles (Liu and Capjak 2006). Different approaches have been tested in the literature. In Dunn (2007) a model has been devised that represents cells by spheres, and the organnelles by ellipsoids. The organelles are randomly placed (but non-overlapping) inside the cell; their size distribution is generated by varying each of the three ellipsoid-axes within ±3 μm about a mean value of 5 μm. It was found that a variation in the organelle volume fraction impacts the phase function at side-scattering angles and the scattering cross section C_{sca}, but not the asymmetry parameter. A variation in the size of the nucleus shows little effect on the phase function when using a simplified homogeneous nucleus model. However, a more realistic inhomogeneous nucleus model yields a

distinct increase in side-scattering intensity when increasing the nucleus size. A model that describes the entire cell by a simple homogeneous sphere model underestimates the phase function at side- and backscattering directions, and it overestimates C_{sca}.

In Zhang et al. (2013) the differences between healthy and carcinogenic cells in diffraction imaging flow cytometry was investigated. The authors considered four different object categories; an inhomogeneous GRS model with nucleus and mitochondria for describing cells with a diameter of 10 μm; a model for irregulary shaped cellular shells; clustered particles consisting of a stack of ellipsoids; and a homogeneous GRS. It was found that the global pattern variation of the diffraction image mainly depends on the object category, and to a lesser extent on the object's inhomogeneous structure. This finding allows one to develop an automated object classification to discriminate among cell debris, intact cells, and different cell morphologies.

In Orlova et al. (2008) the dependence of total and depolarised side-scattering signals on the size and refractive index of cells and cell granules was investigated, as well as on the granule volume fraction. The focus was on granulocytes, which are a subtype of leukocytes. The cells were represented by spheres with multiple, randomly placed spherical inclusions of identical sizes; the model cells had no nucleus. The reference cells had a diameter of $D = 8$ μm, a granule volume fraction $f = 0.1$, and relative refractive indices of 1.015 for the cytoplasm and 1.2 for the granules. The parameters where then varied one at a time, namely, the granule size in the range $d = 0.075$–2 μm, the volume fraction in the range $f = 0.02$–0.3, the cell size in the range $D = 4$–14 μm, and the granules' relative refractive index in the range 1.1–1.2. It was found that the granule size has the strongest impact on total and depolarised side-scattering intensity.

We mentioned that cells are optically soft, which is quite different from many of the particles we encounter in the atmosphere and in solar-system objects. Another distinguishing characteristic in biomedical optics is that cells are often closely spaced, while aerosols and cometary particles are often at distances from each other that are much larger than the wavelength of light. The latter is one of the main prerequisites for the applicability of radiative-transfer theory. As this prerequisite is violated in tissue the use of radiative-transfer theory in biomedical applications is often not well justified. Thus it is important to study the interaction of light with tissue by use of numerically exact methods. However, this task would require enormous computational resources. So far, numerically exact light-scattering computations by dense media of closely-packed cells have only been performed for clusters consisting of a limited number of inhomogeneous cells. In Starosta and Dunn (2010) each cell was represented by a spheroid with axes diameters of 15 and 13 μm. The cells contained a spheroidal nucleus of axes 6 and 5 μm, and two different groups of organelles. The first group consisted of spheroids with axes 1.5 and 0.5 μm, and the second group of spheres with diameter 0.5 μm. The organelles were randomly placed and oriented inside the cell. The cells were then arranged into different cubical grids, the largest of which contained $3 \times 3 \times 4 = 36$ cells, where the incident light propagated in the z direction. Various substructures were

investigated and compared to the full cubical structure. It was found that those substructures that have the complete number of cells in the z direction agreed best with the reference structure. The authors concluded that multiple interaction among cells in the direction of propagation has the dominant impact on the optical properties of the cell cluster.

Acknowledgements We wish to thank Gi Jeong, Konrad Kandler, Osku Kemppinen, Hannakaisa Lindqvist and Olga Muñoz for their contributions to the figures and for helpful comments. M. Kahnert acknowledges funding by the Swedish Research Council (*Vetenskapsrådet*) under contract 621-2011-3346. T. Nousiainen acknowledges funding by the Academy of Finland and the Finnish Funding Agency for Technology and Innovation (TEKES). J. Markkanen was financed through ERC Advanced Grant No 320773 entitled "Scattering and Absorption of ElectroMagnetic waves in ParticuLate media" (SAEMPL).

Bibliography

Ackerman TP, Toon OB (1981) Absorption of visible radiation in atmospheres containing mixtures of absorbing and nonabsorbing particles. Appl Opt 20:3661–3668

Adachi K, Buseck PR (2008) Internally mixed soot, sulfates, and organic matter in aerosol particles from Mexico City. Atmos Chem Phys 8:6469–6481

Adachi K, Chung SH, Friedrich H, Buseck PR (2007) Fractal parameters of individual soot particles determined using electron tomography: implications for optical properties. J Geophys Res 112:D14202. doi:10.1029/2006JD008296

Adachi K, Chung S, Buseck PR (2010) Shapes of soot aerosol particles and implications for their effects on climate. J Geophys Res 115:D15206. doi:10.1029/2009JD012868

Alberts B, Bray D, Lewis J, Raff M, Roberts K, Watson J (1989) Molecular biology of the cell. Garland Publishing, New York

Babenko VA (2010) Electromagnetic scattering in disperse media. Springer, Berlin

Bi L, Yang P, Kattawar G, Kahn R (2009) Single-scattering properties of triaxial ellipsoidal particles for a size parameter range from the Rayleigh to geometric-optics regimes. Appl Opt 48:114–126

Bi L, Yang P, Kattawar G, Kahn R (2010) Modeling optical properties of mineral aerosol particles by using nonsymmetric hexahedra. Appl Opt 49:334–342

Bohren CF (1986) Applicability of effective-medium theories to problems of scattering and absorption by nonhomogeneous atmospheric particles. J Atmos Sci 43:468–475

Bohren CF, Huffman DR (1983) Absorption and scattering of light by small particles. Wiley-VCH, Weinheim

Bond TC, Bergstrom RW (2006) Light absorption by carbonaceous particles: an investigative review. Aerosol Sci Technol 40:27–67

Bond TC, Habib G, Bergstrom RW (2006) Limitations in the enhancement of visible light absorption due to mixing state. J Geophys Res 111:D20211. doi:10.1029/2006JD007315

Bradley J (2004) Interplanetary dust particles. In: Holland H, Turekian K (eds) Treatise on geochemistry. Elsevier, Amsterdam, pp 689–711

Bradley J, Humecki H, Germani M (1992) Combined infrared and analytical electron microscope studies of interplanetary dust particles. Astrophys J 394:643–651

Brownlee D, Tsou P, Aléon J, Alexander C, Araki T, Bajt S, Baratta G, Bastien R, Bland P, Bleuet P, Borg J, Bradley J, Brearley A, Brenker F, Brennan S, Bridges J, Browning N, Brucato J, Bullock E, Burchell M, Busemann H, Butterworth A, Chaussidon M, Cheuvront A, Chi M, Cintala M, Clark B, Clemett S, Cody G, Colangeli L, Cooper G, Cordier P, Daghlian C,

Dai Z, D'Hendecourt L, Djouadi Z, Dominguez G, Duxbury T, Dworkin J, Ebel D, Economou T, Fakra S, Fairey S, Fallon S, Ferrini G, Ferroir T, Fleckenstein H, Floss C, Flynn G, Franchi I, Fries M, Gainsforth Z, Gallien J-P, Genge M, Gilles M, Gillet P, Gilmour J, Glavin D, Gounelle M, Grady M, Graham G, Grant P, Green S, Grossemy F, Grossman L, Grossman J, Guan Y, Hagiya K, Harvey R, Heck P, Herzog G, Hoppe P, Horz F, Huth J, Hutcheon I, Ignatyev K, Ishii H, Ito M, Jacob D, Jacobsen C, Jacobsen S, Jones S, Joswiak D, Jurewicz A, Kearsley A, Keller L, Khodja H, Kilcoyne A, Kissel J, Krot A, Langenhorst F, Lanzirotti A, Le L, Leshin L, Leitner J, Lemelle L, Leroux H, Liu M-C, Luening K, Lyon I, MacPherson G, Marcus M, Marhas K, Marty B, Matrajt G, McKeegan K, Meibom A, Mennella V, Messenger K, Messenger S, Mikouchi T, Mostefaoui S, Nakamura T, Nakano T, Newville M, Nittler L, Ohnishi I, Ohsumi K, Okudaira K, Papanastassiou D, Palma R, Palumbo M, Pepin R, Perkins D, Perronnet M, Pianetta P, Rao W, Rietmeijer F, Robert F, Rost D, Rotundi A, Ryan R, Sandford S, Schwandt C, See T, Schlutter D, Sheffield-Parker J, Simionovici A, Simon S, Sitnitsky I, Snead C, Spencer M, Stadermann F, Steele A, Stephan T, Stroud R, Susini J, Sutton S, Suzuki Y, Taheri M, Taylor S, Teslich N, Tomeoka K, Tomioka N, Toppani A, Trigo-Rodrguez J, Troadec D, Tsuchiyama A, Tuzzolino A, Tyliszczak T, Uesugi K, Velbel M, Vellenga J, Vicenzi E, Vincze L, Warren J, Weber I, Weisberg M, Westphal A, Wirick S, Wooden D, Wopenka B, Wozniakiewicz P, Wright I, Yabuta H, Yano H, Young E, Zare R, Zega T, Ziegler K, Zimmerman L, Zinner E, Zolensky M (2006) Comet 81P/Wild 2 under a microscope. Science 314:1711–1716

Chang H, Charalampopoulos TT (1990) Determination of the wavelength dependence of refractive indices of flame soot. Proc R Soc Lond A 430:577–591

Cheng T, Wu Y, Chen H (2014) Effects of morphology on the radiative properties of internally mixed light absorbing carbon aerosols with different aging status. Opt Express 22:15904–15917

Chýlek P, Videen G, Geldart DJW, Dobbie JS, Tso HCW (2000) Effective medium approximations for heterogeneous particles. In: Mishchenko MI, Hovenier JW, Travis LD (eds) Light scattering by nonspherical particles. Academic Press, San Diego, pp 274–308

Colbeck I, Appleby L, Hardman EJ, Harrison RM (1990) The optical properties and morphology of cloud-processed carbonaceous smoke. J Aerosol Sci 21:527–538

Coz E, Leck C (2011) Morphology and state of mixture of atmospheric soot aggregates during the winter season over southern Asia—A quantitative approach. Tellus 63B:107116

Dabrowska D, Muñoz O, Moreno F, Nousiainen T, Zubko E (2012) Effect of the orientation of the optic axis on simulated scattering matrix elements of small birefringent particles. Opt Lett 37:3252–3254

Draine BT (2000) The discrete dipole approximation for light scattering by irregular targets. In: Mishchenko MI, Hovenier JW, Travis LD (eds) Light scattering by nonspherical particles. Academic Press, San Diego, pp 131–144

Draine BT, Flatau PJ (1994) Discrete-dipole approximation for scattering calculations. J Opt Soc Am A 11:1491–1499

Dubovik O, Sinyuk A, Lapyonok T, Holben BN, Mishchenko MI, Yang P, Eck TF, Volten H, Muñoz O, Veihelmann B, van der Zande WJ, Leon J-F, Sorokin M, Slutsker I (2006) Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust. J Geophys Res 111:D11208

Dunn AK (2007) Modelling of light scattering from inhomogeneous biological cells. In: Hoekstra A, Maltsev VP, Videen G (eds) Optics of biological particles. Springer, Dordrecht, pp 19–29

Durant A, Bonadonna C, Horwell C (2010) Atmospheric and environmental impacts of volcanic particulates. Elements 6:235–240

Eddington AS (1926) The internal constitution of the stars. Cambridge University Press, London

Fuller KA (1995) Scattering and absorption cross sections of compounded spheres III. Spheres containing arbitrarily located spherical inhomogeneities. J Opt Soc Am A 12:893–904

Fuller KA, Malm WC, Kreidenweis SM (1999) Effects of mixing on extinction by carbonaceous particles. J Geophys Res 104:15941–15954

Gasteiger J, Wiegner M, Groß S, Freudenthaler V, Toledano C, Tesche M, Kandler K (2011) Modelling lidar-relevant optical properties of complex mineral dust aerosols. Tellus 63B:725–741

Greenberg J (1982) What are comets made of? A model based on interstellar dust. In: Wilkening L (ed) comets. University of Arizona Press, Tuscon, pp 131–163

Greenberg J, Hage J (1990) From interstellar dust to comets—A unification of observational constraints. Astrophys J 361:260–274

Gupta R, Vaidya D, Bobbie J, Chýlek P (2006) Scattering properties and composition of cometary dust. Astrophys Space Sci 301:21–31

Hanner M, Bradley J (2004) Effective medium approximations for heterogeneous particles. In: Festou M, Keller H, Weaver H (eds) Comets II. University of Arizona Press, Tucson, pp 555–564

Hess M, Koepke P, Schult I (1998) Optical properties of aerosols and clouds: The software package OPAC. Bull Am Met Soc 79:831–844

Ishii H, Bradley J, Rong Dai Z, Chi M, Kearsley A, Burchell M, Browning N, Molster F (2008) Comparison of comet 81P/Wild 2 dust with interplanetary dust from comets. Science 319: 447–450

Ishimoto H, Zaizen Y, Uchiyama A, Masuda K, Mano Y (2010) Shape modeling of mineral dust particles for light-scattering calculations using the spatial poissonvoronoi tessellation. J Quant Spectrosc Radiat Transfer 111:2434–2443

Jeong G, Nousiainen T (2014) TEM analysis of the internal structures and mineralogy of asian dust particles and the implications for optical modeling. Atmos Chem Phys 14:7233–7254

Jessberger E, Christoforidis A, Kissel J (1988) Aspects of the major element composition of Halley's dust. Nature 332:691–695

Jones AR (2006) Light scattering in combustion. In: Kokhanovsky A (ed) Light scattering reviews. Springer, Berlin, pp 393–444

Kahnert M (2010a) Modelling the optical and radiative properties of freshly emitted light absorbing carbon within an atmospheric chemical transport model. Atmos Chem Phys 10:1403–1416

Kahnert M (2010b) Numerically exact computation of the optical properties of light absorbing carbon aggregates for wavelength of 200 nm 12.2 μm. Atmos Chem Phys 10:8319–8329

Kahnert M (2010c) On the discrepancy between modelled and measured mass absorption cross sections of light absorbing carbon aerosols. Aerosol Sci Technol 44:453–460

Kahnert M (2015) Modelling radiometric properties of inhomogeneous mineral dust particles: Applicability and limitations of effective medium theories. J Quant Spectrosc Radiat Transfer 152:16–27

Kahnert M, Devasthale A (2011) Black carbon fractal morphology and short-wave radiative impact: a modelling study. Atmos Chem Phys 11:11745–11759

Kahnert M, Nousiainen T, Lindqvist H, Ebert M (2012) Optical properties of light absorbing carbon aggregates mixed with sulfate: assessment of different model geometries for climate forcing calculations. Opt Express 20:10042–10058

Kahnert M, Nousiainen T, Lindqvist H (2013) Models for integrated and differential scattering optical properties of encapsulated light absorbing carbon aggregates. Opt Express 21:7974–7992

Kahnert M, Nousiainen T, Lindqvist H (2014) Review: model particles in atmospheric optics. J Quant Spectrosc Radiat Transfer 146:41–58

Kalashnikova O, Kahn R, Sokolik I, Li W-H (2005) Ability of multiangle remote sensing observations to identify and distinguish mineral dust types: optical models and retrievals of optically thick plumes. J Geophys Res 110:D18S14

Kandler K, Schütz L, Deutscher C, Hofmann H, Jäckel S, Knippertz P, Lieke K, Massling A, Schladitz A, Weinzierl B, Zorn S, Ebert M, Jaenicke R, Petzold A, Weinbruch S (2009) Size distribution, mass concentration, chemical and mineralogical composition, and derived optical parameters of the boundary layer aerosol at Tinfou, Morocco, during SAMUM 2006. Tellus 61B:32–50

Keller L, Messenger S, Bradley J (2000) Analysis of a deuterium-rich interplanetary dust particle (IDP) and implications for presolar material in IDPs. J Geophys Res Space Phys 105:2156–2202

Kemppinen O, Nousiainen T, Jeong GY (2015) Effects of dust particle internal structure on light scattering. Atmos Chem Phys Discuss 15:20349–20394

Kimura H, Kolokolova L, Mann I (2006) Light scattering by cometary dust numerically simulated with aggregate particles consisting of identical spheres. Astron Astrophys 449:1243–1254

Kiselev N, Velichco S, Jockers K, Rosenbush V, Kikuchi S (2006) Database of comet polarimetry. EAR-C-COMPIL-5-DB-COMET-POLARIMETRY-V1.0. NASA Planetary Data System

Kocifaj M, Videen G (2008) Optical behavior of composite carbonaceous aerosols: DDA and EMT approaches. J Quant Spectrosc Radiat Transfer 109:1404–1416

Kokhanovsky AA (2004) Light scattering media optics: problems and solutions. Springer, Berlin

Kolokolova L, Hanner M, Levasseur-Regourd A-C, Gustafson B (2004) Physical properties of cometary dust from light scattering and thermal emission. In: Festou M, Keller U, Weaver H (eds) Comets II. University of Arizona Press, Tucson, pp 577–604

Krotkov N, Flittner D, Krueger A, Kostinski A, Riley C, Rose W, Torres O (1999) Effect of particle non-sphericity on satellite monitoring of drifting volcanic ash clouds. J Quant Spectrosc Radiat Transfer 63:613–630

Kylling A, Kahnert M, Lindqvist H, Nousiainen T (2014) Volcanic ash infrared signature: porous non-spherical ash particle shapes compared to homogeneous spherical ash particles. Atmos Meas Tech 7:919–929

Lasue J, Levasseur-Regourd A-C (2006) Porous irregular aggregates of sub-micron sized grains to reproduce cometary dust light scattering observations. J Quant Spectrosc Radiat Transfer 100:220–236

Lasue J, Levasseur-Regourd A-C, Hadamcik E, Alcouffe G (2009) Cometary dust properties retrieved from polarization observations: application to C/1995 O1 Hale-Bopp and 1P/Halley. Icarus 199:129–144

Lawler M, Brownlee D (1992) CHON as a component of dust from comet Halley. Nature 359:810812

Levasseur-Regourd A-C, Mukai T, Lasue J, Okada Y (2007) Physical properties of cometary and interplanetary dust. Planet Space Sci 55:1010–1020

Lieke K, Kristensen T, Korsholm U, Sørensen J, Kandler K, Weinbruch S, Ceburnis D, Ovadnevaite J, ODowd C, Bilde M (2013) Characterization of volcanic ash from the 2011 grímsvötn eruption by means of single-particle analysis. Atmos Environ 79:411–420

Lindqvist H, Muinonen K, Nousiainen T (2009) Light scattering by coated Gaussian and aggregate particles. J Quant Spectrosc Radiat Transfer 110:1398–1410

Lindqvist H, Nousiainen T, Zubko E, Muñoz O (2011) Optical modeling of vesicular volcanic ash particles. J Quant Spectrosc Radiat Transfer 112:1871–1880

Lindqvist H, Jokinen O, Kandler K, Scheuvens D, Nousiainen T (2014) Single scattering by realistic, inhomogeneous mineral dust particles with stereogrammetric shapes. Atmos Chem Phys 14:143–157

Liu C, Capjak CE (2006) Effects of cellular fine structure on scattered light pattern. IEEE Trans Nanobiosci 5:76–82

Liu L, Mishchenko MI (2007) Scattering and radiative properties of complex soot and soot-containing aggregate particles. J Quant Spectrosc Radiat Transfer 106:262–273

Liu L, Mishchenko MI, Arnott WP (2008) A study of radiative properties of fractal soot aggregates using the superposition *T*-matrix method. J Quant Spectrosc Radiat Transfer 109:2656–2663

Liu C, Panetta RL, Yang P, Macke A, Baran AJ (2012) Modeling the scattering properties of mineral aerosols using concave fractal polyhedra. Appl Opt 52:640–652

Lumme K, Penttilä A (2011) Model of light scattering by dust particles in the solar system: applications to cometary comae and planetary regoliths. J Quant Spectrosc Radiat Transfer 112:1658–1670

Mackowski DW, Mishchenko MI (1996) Calculation of the T matrix and the scattering matrix for ensembles of spheres. J Opt Soc Am A 13:2266–2278

Mackowski DW, Mishchenko MI (2011) A multiple sphere T-matrix Fortran code for use on parallel computer clusters. J Quant Spectrosc Radiat Transfer 112:2182–2192
Markkanen J, Penttilä A, Peltoniemi J, Muinonen K (2015) Model for light-scattering by cometary dust. Planet Space Sci 118:164–172
McFiggans G, Artaxo P, Baltensperger U, Coe H, Facchini MC, Feingold G, Fuzzi S, Gysel M, Laaksonen A, Lohmann U, Mentel TF, Murphy DM, O'Dowd CD, Snider JR, Weingartner E (2006) The effect of physical and chemical aerosol properties on warm cloud droplet activation. Atmos Chem Phys 6:2593–2649
Merikallio S, Muñoz O, Sundström A-M, Virtanen T, Horttanainen M, de Leeuw G, Nousiainen T (2015) Optical modeling of volcanic ash particles using ellipsoids. J Geophys Res Atmos 120:4102–4116
Mishchenko MI, Lacis AA, Carlson BE, Travis LD (1995) Nonsphericity of dust-like tropospheric aerosols: implications for aerosol remote sensing and climate modeling. Geophys Res Lett 22:1077–1080
Mishchenko MI, Liu L, Travis LD, Lacis AA (2004) Scattering and radiative properties of semi-external versus external mixtures of different aerosol types. J Quant Spectrosc Radiat Transfer 88:139–147
Mishchenko MI, Dlugach ZM, Zakharova NT (2014) Direct demonstration of the concept of unrestricted effective-medium approximation. Opt Lett 39:3935–3938
Mukai S, Mukai T, Kikuchi S (1991) Scattering properties of cometary dust based on polarimetric data. In: Levasseur-Regourd A-C, Hasegawa H (eds) Origin and evolution of interplanetary dust. Springer, Berlin, pp 249–252
Muñoz O, Volten H, Hovenier JW, Veihelmann B, van der Zande WJ, Waters LBFM, Rose WI (2004) Scattering matrices of volcanic ash particles of Mount St. Helens, Redoubt, and Mount Spurr volcanoes. J Geophys Res 109
Niemeier U, Timmreck C, Graf H-F, Kinne S, Rast S, Self S (2009) Initial fate of fine ash and sulfur from large volcanic eruptions. Atmos Chem Phys 9:9043–9057
Nousiainen T, Kandler K (2015) Light scattering by atmospheric mineral dust particles. In: Kokhanovsky A (ed) Light scattering reviews, vol 9, Chap 1. Springer, Berlin, 430 pp
Nousiainen T, Vermeulen K (2003) Comparison of measured single-scattering matrix of feldspar particles with T-matrix simulations using spheroids. J Quant Spectrosc Radiat Transfer 79–80:1031–1042
Nousiainen T, Muinonen K, Räisänen P (2003) Scattering of light by Saharan particles in a modified ray-optics approximation. J Geophys Res 108. doi:10.1029/2001JD001277
Nousiainen T, Kahnert M, Veihelmann B (2006) Light scattering modeling of small feldspar aerosol particles using polyhedral prisms and spheroids. J Quant Spectrosc Radiat Transfer 101:471–487
Nousiainen T, Zubko E, Niemi JV, Kupiainen K, Lehtinen M, Muinonen K, Videen G (2009) Single-scattering modeling of thin, birefringent mineral-dust flakes using the discrete-dipole approximation. J Geophys Res 114:D07207. doi:10.1029/2008JD011564
Nousiainen T, Kahnert M, Lindqvist H (2011) Can particle shape information be retrieved from light-scattering observations using spheroidal model particles? J. Quant Spectrosc Radiat Transfer 112:2213–2225
Orlova DY, Yurkin MA, Hoekstra AG, Maltsev VP (2008) Light scattering by neutrophils: model, simulation, and experiment. J Biomed Opt 13:054057
Petrova E, Jockers K, Kiselev N (2000) Light scattering by aggregates with sizes comparable to the wavelength: An application to cometary dust. Icarus 148:526–536
Purcell EM, Pennypacker CR (1973) Scattering and absorption of light by nonspherical dielectric grains. Astrophys J 186:705–714
Räisänen P, Haapanala P, Chung CE, Kahnert M, Makkonen R, Tonttila J, Nousiainen T (2013) Impact of dust particle non-sphericity on climate simulations. Q J R Meteorol Soc 139:2222–2232

Ramachandran G, Reist PC (1995) Characterization of morphological changes in agglomerates subject to condensation and evaporation using multiple fractal dimensions. Aerosol Sci Technol 23:431–442

Sandford S, Aléon J, Alexander C, Araki T, Bajt S, Baratta G, Borg J, Bradley J, Brownlee D, Brucato J, Burchell M, Busemann H, Butterworth A, Clemett S, Cody G, Colangeli L, Cooper G, D'Hendecourt L, Djouadi Z, Dworkin J, Ferrini G, Fleckenstein H, Flynn G, Franchi I, Fries M, Gilles M, Glavin D, Gounelle M, Grossemy F, Jacobsen C, Keller L, Kilcoyne A, Leitner J, Matrajt G, Meibom A, Mennella V, Mostefaoui S, Nittler L, Palumbo M, Papanastassiou D, Robert F, Rotundi A, Snead C, Spencer M, Stadermann F, Steele A, Stephan T, Tsou P, Tyliszczak T, Westphal A, Wirick S, Wopenka B, Yabuta H, Zare R, Zolensky M (2006) Organics captured from comet 81P/Wild 2 by the Stardust spacecraft. Science 314:1720–1724

Scarnato B, Vahidinia S, Richard DT, Kirchstetter TW (2013) Effects of internal mixing and aggregate morphology on optical properties of black carbon using a discrete dipole approximation model. Atmos Chem Phys 13:5089–5101

Scheuvens D, Kandler K, Küpper M, Lieke K, Zorn S, Ebert M, Schütz L, Weinbruch S (2011) Indiviual-particle analysis of airborne dust samples collected over Morocco in 2006 during SAMUM 1. Tellus 63B:512–530

Schmidt K, Yurkin M, Kahnert M (2012) A case study on the reciprocity in light scattering computations. Opt Express 20:23253–23274

Schnaiter M, Horvath H, Möhler O, Naumann K-H, Saathoff H, Schöck OW (2003) UV-VIS-NIR spectral optical properties of soot and soot-containing aerosols. J Aerosol Sci 34:1421–1444

Schnaiter M, Linke C, Moehler O, Naumann K-H, Saathoff H, Wagner R, Schurath U, Wehner B (2005) Absorption amplification of black carbon internally mixed with secondary organic aerosol. J Geophys Res 110:D19204. doi:10.1029/2005JD006046

Schumann U, Weinzierl B, Reitebuch O, co authors (2011) Airborne observations of the Eyjafjalla volcano ash cloud over Europe during air space closure in April and May 2010. Atmos Chem Phys 11:2245–2279

Sihvola A (1999) Electromagnetic mixing formulas and applications. Institution of Electrical Engineers, London

Skorupski K, Mroczka J, Riefler N, Oltmann H, Will S, Wriedt T (2013) Impact of morphological parameters onto simulated light scattering patterns. J Quant Spectrosc Radiat Transfer 119:53–66

Starosta MS, Dunn AK (2010) Far-field superposition method for three-dimensional computation of light scattering from multiple cells. J Biomed Opt 15:055006

Stevenson J, Millington S, Beckett F, Swindles G, Thordarson T (2015) Big grains go far: reconciling tephrochronology with atmospheric measurements of volcanic ash. Atmos Meas Tech D 8:65–120

van Poppel LH, Friedrich H, Spinsby J, Chung SH, Seinfeld JH, Buseck PR (2005) Electron tomography of nanoparticle clusters: Implications for atmospheric lifetimes and radiative forcing of soot. Geophys Res Lett 32:L24811. doi:10.1029/2005GL024461

Veihelmann B, Nousiainen T, Kahnert M, van der Zande WJ (2006) Light scattering by small feldspar particles simulated using the gaussian random sphere geometry. J Quant Spectrosc Radiat Transfer 100(1–3):393–405

Videen G, Chýlek P (1998) Scattering by a composite sphere with an absorbing inclusion and effective medium approximations. Opt Commun 158:1–6

Videen G, Zubko E, Sun W, Shkuratov Y, Yuffa A (2015) Mixing rules and morphology dependence of the scatterer. J Quant Spectrosc Radiat Transfer 150:68–75

Vilaplana R, Moreno F, Molina A (2006) Study of the sensitivity of size-averaged scattering matrix elements of nonspherical particles to changes in shape, porosity and refractive index. J Quant Spectrosc Radiat Transfer 100:415–428

Wang Z, Zhang H, Jing X, Wei X (2013) Effect of non-spherical dust aerosol on its direct radiative forcing. Atmos Res 120–121:112–126

Whittet DCB (2003) Dust in the galactic environment. Institute of Physics Publishing, Bristol

Worringen A, Ebert M, Trautmann T, Weinbruch S, Helas G (2008) Optical properties of internally mixed ammonium sulfate and soot particles—A study of individual aerosol particles and ambient aerosol populations. Appl Opt 47:3835–3845

Wriedt T, Wilkens J, Hellmers J (2010) Differentiating between sintered and non-sintered aggregates. In: Muinonen K, Penttilä A, Lindqvist H, Nousiainen T, Videen G (eds) Electromagnetic and Light Stattering XII. University of Helsinki, Helsinki, pp 322–325

Wu Y, Cheng T, Zheng L, Chen H (2015a) Effect of morphology on the optical properties of soot aggregated with spheroidal monomers. J Quant Spectrosc Radiat Transfer

Wu Y, Cheng T, Zheng L, Chen H (2015b) A study of optical properties of soot aggregates composed of poly-disperse monomers using the superposition T-matrix method. Aerosol Sci Technol 49:941–949

Wu Y, Cheng T, Zheng L, Chen H, Xu H (2015c) Single scattering properties of semi-embedded soot morphologies with intersecting and non-intersecting surfaces of absorbing spheres and non-absorbing host. J Quant Spectrosc Radiat Transfer 157:1–13

Xing Z, Hanner M (1997) Light scattering by aggregate particles. Astron Astrophys 324:805–820

Xu Y, Gustafson BÅS (2001) A generalized multiparticle Mie solution: further experimental verification. J Quant Spectrosc Radiat Transfer 70:395–419

Yanamandra-Fisher P, Hanner M (1999) Optical properties of nonspherical particles of size comparable to the wavelength of light: application to comet dust. Icarus 139:388–389

Yurkin MA, Hoekstra AG (2011) The discrete-dipole-approximation code ADDA: capabilities and known limitations. J Quant Spectrosc Radiat Transfer 112:2234–2247

Zhang J, Feng Y, Moran MS, Lu JQ, Yang LV, Sa Y, Zhang N, Dong L, Hu X-H (2013) Analysis of cellular objects through diffraction images acquired by flow cytometry. Opt Express 21:2481924828

Zhao Y, Ma L (2009) Assessment of two fractal scattering models for the prediction of the optical characteristics of soot aggregates. J Quant Spectrosc Radiat Transfer 110:315–322

Zolensky M, Zega T, Yano H, Wirick S, Westphal A, Weisberg M, Weber I, Warren J, Velbel M, Tsuchiyama A, Tsou P, Toppani A, Tomioka N, Tomeoka K, Teslich N, Taheri M, Susini J, Stroud R, Stephan T, Stadermann F, Snead C, Simon S, Simionovici A, See T, Robert F, Rietmeijer F, Rao W, Perronnet M, Papanastassiou D, Okudaira K, Ohsumi K, Ohnishi I, Nakamura-Messenger K, Nakamura T, Mostefaoui S, Mikouchi T, Meibom A, Matrajt G, Marcus M, Leroux H, Lemelle L, Le L, Lanzirotti A, Langenhorst F, Krot A, Keller L, Kearsley A, Joswiak D, Jacob D, Ishii H, Harvey R, Hagiya K, Grossman L, Grossman J, Graham G, Gounelle M, Gillet P, Genge M, Flynn G, Ferroir T, Fallon S, Ebel D, Rong Dai Z, Cordier P, Clark B, Chi M, Butterworth A, Brownlee D, Bridges J, Brennan S, Brearley A, Bradley J, Bleuet P, Bland P, Bastien R (2006) Mineralogy and petrology of comet 81P/Wild 2 nucleus samples. Science 314(5806):1735–1739

Zubko E, Furusho R, Kawabata K, Yamamoto T, Muinonen K, Videen G (2011) Interpretation of photo-polarimetric observations of comet 17P/Holmes. J Quant Spectrosc Radiat Transfer 112:1848–1863

Zubko E, Muinonen K, Muñoz O, Nousiainen T, Shkuratov Y, Sun W, Videen G (2013) Light scattering by feldspar particles: comparison of model agglomerate debris particles with laboratory samples. J Quant Spectrosc Radiat Transfer 131:175–187

Zubko E, Muinonen K, Videen G, Kiselev N (2014) Dust in comet C/1975 V1 (West). Mon Not R Astron Soc 440:2928–2943

Some Wave-Theoretic Problems in Radially Inhomogeneous Media

Umaporn Nuntaplook, John A. Adam and Michael A. Pohrivchak

1 Introduction

With recent renewed interest in the classical topics of both acoustic and electromagnetism as applied to nanotechnology, transformation optics, fiber optics, metamaterials with negative refractive indices, cloaking and invisibility, the topic of time-independent scattering theory in quantum mechanics is becoming a useful field to reexamine in the above contexts. One of the key areas of electromagnetic theory—scattering of plane electromagnetic waves—is highly dependent on the properties of the refractive indices in the various media. Furthermore, the refractive index of a medium and the potential in quantum scattering theory are intimately related. In many cases, understanding scattering in radially symmetric media is necessary to gain insight into scattering in more complex media. Meeting the challenge of variable refractive indices and possibly complicated boundary conditions, therefore ,not only requires accurate and efficient numerical methods, but where possible, analytic solutions to the governing radial wave equations. Despite common emphasis on constant refractive indices, the most interesting (and increasingly useful) cases are those with radially inhomogeneous refractive index

U. Nuntaplook (✉)
Department of Mathematics, Faculty of Science, Mahidol University, Bangkok, Thailand
e-mail: ununtaplook@gmail.com

J.A. Adam · M.A. Pohrivchak
Department of Mathematics and Statistics, Old Dominion University, Norfolk, VA, USA
e-mail: jadam@odu.edu

M.A. Pohrivchak
e-mail: mporkchop19@gmail.com

M.A. Pohrivchak
Naval Research Laboratory, Washington, DC, USA

© Springer-Verlag Berlin Heidelberg 2016

A. Kokhanovsky (ed.), *Light Scattering Reviews, Volume 11*,
Springer Praxis Books, DOI 10.1007/978-3-662-49538-4_7

profiles. Valuable insights can be gained by treating an inhomogeneous medium as a piecewise constant one, and that is the primary focus of this chapter.

The question may be asked as to why, with all the powerful computational facilities at our disposal, analytic solutions for radially inhomogeneous media [even ray-theoretic solutions; see Pohrivchak (2014)] should be of more than historical interest today. Furthermore, what are some of the potential applications for such solutions? There are several different contexts to be considered in answering these questions. Idealized planetary atmospheres (for example) may be treated as spheres surrounded by a large spherical shell with a radially inhomogeneous dielectric constant. In addition, the plasma sheath surrounding a satellite reentering the Earth's atmosphere may possess something close to radial inhomogeneity (Brockman 1974). In the biological realm, the scattering patterns produced from bacteria may in some cases be modeled by radially inhomogeneous density distributions (Fenn and Oser 1965; Brunsting and Mullaney 1972; Kitchen and Zaneveld 1992). Dielectric lens design in radar and optical scattering (Cornbleet 1984), nonuniform optical fibers (Adams 1981), etc. In all the above examples, exact solutions (and their limiting cases) can be of use, especially since expressions for the amplitude and phase of the associated electromagnetic field can be directly compared with the corresponding behavior as predicted by geometrical optics. As a result, corrections to the latter can be formulated; such corrections are well-known, for example, when rays pass through focal points or caustic curves (Alexopoulos 1971, 1972, 1974).

More recently, with the advent of research in nanotechnology and metamaterials, transformation optics and cloaking (Pendry et al. 2006; Leonhardt 2006), analytic solutions are as valuable as they ever were (despite advances in numerical studies of such solutions), if not more so. Indeed, as Leonhardt and Philbin (2010) point out, "In a metamaterial, structures smaller than the wavelength of light control the optical properties of the material. Their shapes and sizes matter more than their chemistry…Thanks to advances in modern nanotechnology and the science behind it, engineers can now make carefully controlled subwavelength structures with designs based on accurate theoretical predictions…"

The wave-theoretic aspects of this chapter are based on the solution of Maxwell's equations for scattering of plane electromagnetic waves from a dielectric (or "transparent") sphere in terms of the related Helmholtz equation. There is a connection with the time-independent Schrödinger equation in the following sense: the time-dependent Schrödinger equation is identical to the wave equation for the scalar radiation potential for TE-polarized electromagnetic waves. In regions where the refractive index is constant, it is also identical to the scalar radiation potential for TM-polarized electromagnetic waves, but with different boundary conditions than for the TE case. We examine scattering of the TE mode from a piecewise-uniform radial inhomogeneity embedded in an external medium (as opposed to an off-axis inclusion), and are a natural generalization of previously known results. The corresponding theory for the TM mode is also developed, and the well-known connection with morphology-dependent resonances (MDRs) in these contexts is noted.

2 Wave Theory for Piecewise-Homogeneous Spheres

In (Adam 2014) many of the existing analytical solutions for the radial wave equations were reviewed in the context of plane wave electromagnetic scattering (see also Adam 2013). As might be expected, the mathematics is rather complicated (involving several special functions such as Bessel, Whittaker, and hypergeometric functions). Valuable though this approach can be, a somewhat different approach to such scattering problems is to use piecewise inhomogeneous spheres. Nuntaplook (2013) used the technique developed by Johnson (1993) based on solving the differential equations for the radial Debye potential, but for a two-layer inhomogeneous sphere. embedded in a uniform medium (see also Nuntaplook and Adam 2014). In this section, we summarize Johnson's method and results for a uniform sphere embedded in the same external medium. The radius of the particle is denoted by a (which of course can be scaled to unity), the center is at the origin of the coordinate system, and the refractive index profile is defined as $n(r)$, i.e., it is considered a function of the radial coordinate r. For the external region outside the sphere, $r > a$, the refractive index, $n(r)$ is equal to one. The wavenumber is $k = 2\pi/\lambda$, where λ is the wavelength outside the sphere. We assume that the particle is nonmagnetic. The complex time-dependence of the electric field is assumed to be harmonic, i.e., $\propto \exp(-i\omega t)$.

We consider the following vector Helmholtz equation for the electric field vector $\mathbf{E}$:

$$\nabla\times\nabla\times\mathbf{E} - k^2 n^2(r)\mathbf{E} = 0. \tag{1}$$

It can be verified that the following vector wave functions are solutions of this equation:

$$\mathbf{M}(r,\theta,\phi) = \nabla \times [\Psi(r,\theta,\phi)\mathbf{r}], \tag{2}$$

and

$$\mathbf{N}(r,\theta,\phi) = \frac{1}{kn^2(r)}\nabla \times \nabla \times [\Phi(r,\theta,\phi)\mathbf{r}], \tag{3}$$

where r, θ, and ϕ are spherical coordinates, $\mathbf{r}$ is the radius vector, and Ψ and Φ are scalar functions that satisfy the following differential equations: (Wyatt 1962)

$$\nabla^2 \Psi + k^2 n^2(r)\Psi = 0, \tag{4}$$

and

$$\nabla^2\Phi - \frac{1}{n^2(r)}\frac{\mathrm{d}n^2(r)}{\mathrm{d}r}\frac{\mathrm{d}\Phi}{\mathrm{d}r} + k^2 n^2(r)\Phi = 0. \tag{5}$$

By applying the method of separation of variables to these equations in spherical coordinates, we can generate the following sets of solutions:

$$\Psi_{l,m}(r,\theta,\phi) = \frac{1}{kr} S_l(r) P_l^m(\cos\,\theta)\exp(im\phi), \tag{6}$$

$$\Phi_{l,m}(r,\theta,\phi) = \frac{1}{kr} T_l(r) P_l^m(\cos\theta)\exp(im\phi), \tag{7}$$

where the $P_l^m(\cos\,\theta)$ are associated Legendre polynomials. The functions $S_l(r)$ and $T_l(r)$ are the radial Debye potentials, which satisfy the following second-order differential equations:

$$\frac{\mathrm{d}^2 S_l(r)}{\mathrm{d}r^2} + \left[k^2 n^2(r) - \frac{l(l+1)}{r^2}\right] S_l(r) = 0, \tag{8}$$

$$\frac{\mathrm{d}^2 T_l(r)}{\mathrm{d}r^2} - \left(\frac{2}{n(r)}\frac{\mathrm{d}n(r)}{\mathrm{d}r}\right)\frac{\mathrm{d}T_l(r)}{\mathrm{d}r} + \left[k^2 n^2(r) - \frac{l(l+1)}{r^2}\right] T_l(r) = 0. \tag{9}$$

In regions where the refractive index is a constant value n, the Eqs. (8) and (9) have the same form. Substituting the functions $\Psi_{l,m}(r,\theta,\phi)$ and $\Phi_{l,m}(r,\theta,\phi)$ into (8) and (9), respectively, provides the set of vector wave functions $\mathbf{M}_{l,m}$ and $\mathbf{N}_{l,m}$ that are solutions to Eq. (1). The **M** fields are called transverse electric (TE) modes, and the **N** fields are transverse magnetic (TM) modes. Therefore, the $S_l(r)$ functions are associated with TE fields, and the $T_l(r)$ are associated with TM fields. In the internal region for a constant refractive index n, $0 \le r \le a$:

$$S_l(r) = T_l(r) = \psi_l(nkr), \tag{10}$$

where $\psi_l(nkr)$ is a Riccati–Bessel function (defined below) and the boundary conditions are $S_l(0) = 0$ and $T_l(0) = 0$, which guarantee that the electric fields and magnetic fields are finite at the origin.

In the region outside the particle the general solutions to (8) and (9) are linear combinations of the Riccati–Bessel functions. The electric field in this region consists of the incident wave and an outgoing scattered wave. The Riccati–Bessel functions are defined as

$$\psi_l(x) = x j_l(x), \tag{11}$$

$$\chi_l(x) = x y_l(x), \tag{12}$$

where $j_l(x)$ and $y_l(x)$ are spherical Bessel function of the first and second kinds, respectively. In the external region: $r > a$.

$$S_l(r) = B_l[\chi_l(kr) + \beta_l \psi_l(kr)], \tag{13}$$

$$T_l(r) = A_l[\chi_l(kr) + \alpha_l \psi_l(kr)], \tag{14}$$

where $n(r) = 1$, α_l, β_l, A_l, and B_l are constants. We use the log-derivative formalism for matching the internal and external solutions where the refractive index is discontinuous, such as at the surface of the sphere and at the boundaries between the layers of the two-layer sphere case in the next section. The modified log-derivative functions of $S_l(r)$ and $T_l(r)$ are defined as

$$U_l(r) = \frac{1}{k}\left(\frac{S_l'(r)}{S_l(r)}\right), \tag{15}$$

$$V_l(r) = \frac{1}{kn^2(r)}\left(\frac{T_l'(r)}{T_l(r)}\right), \tag{16}$$

where a prime denotes the derivative with respect to the argument of the function. Both of these functions are continuous at all points. We first consider the continuity at the surface $r = a$ by matching the internal solution from (10) with the external solution from (13) and (14) to obtain

$$\psi_l(nka) = B_l[\chi_l(ka) + \beta_l \psi_l(ka)], \tag{17}$$

$$\psi_l(nka) = A_l[\chi_l(ka) + \alpha_l \psi_l(ka)]. \tag{18}$$

Hence,

$$B_l = \frac{\psi_l(nka)}{\chi_l(ka) + \beta_l \psi_l(ka)}, \tag{19}$$

$$A_l = \frac{\psi_l(nka)}{\chi_l(ka) + \alpha_l \psi_l n(ka)}. \tag{20}$$

As noted by Johnson (1993), were the barrier to extend to infinity, the wave function would decay to zero and the 'quasi-bound state' would become a true bound state. In this limit only the decreasing function $\chi_l(kr)$ is permitted in the barrier region. This requires the coefficient of $\psi_l(kr)$ be zero at the location of a TE resonance (i.e., $\beta_l = 0$; for the TM mode $\alpha_l = 0$). While these conditions (on the basis of the limiting case of an infinite barrier) appear to be only sufficient for a finite barrier, from Hill and Benner (1988) they are seen to be equivalent to those defining the existence of a MDR, e.g., from Eq. (24) below, $\mathrm{Im}\, b_l = 0$ (or $\mathrm{Im}\, a_l = 0$ for the TM mode). For $\beta_l = 0$ and $\alpha_l = 0$ at resonance it follows that

$$A_l = B_l = \frac{\psi_l(nka)}{\chi_l(ka)}. \tag{21}$$

Now, we substitute the external solutions defined by (13) and (14) into the modified logarithmic derivatives given by (15) and (16) and evaluate them at the surface of the sphere $r = a$. The standard approach in the literature is to use the continuity of the functions $U_l(r)$ and $V_l(r)$ across the boundary to solve for α_l and β_l. The results are

$$\beta_l = -\frac{\psi_l(nx)\chi_l'(x) - n\psi_l'(nx)\chi_l(x)}{\psi_l(nx)\psi_l'(x) - n\psi_l'(nx)\psi_l(x)}, \tag{22}$$

$$\alpha_l = -\frac{\psi_l'(nx)\chi_l(x) - n\psi_l(nx)\chi_l'(x)}{\psi_l'(nx)\psi_l(x) - n\psi_l(nx)\psi_l'(x)}, \tag{23}$$

where $x = ka$ is the size parameter. The coefficients α_l and β_l are related to the a_l and b_l coefficients of Mie theory by the following formulas:

$$b_l = \frac{1}{1 - i\beta_l}, \tag{24}$$

$$a_l = \frac{1}{1 - i\alpha_l}. \tag{25}$$

The a_l and b_l coefficients defined here are the same as the coefficients defined by Bohren and Huffman (1983) and are the complex conjugate of the coefficients defined by van de Hulst (1981) and by Kerker (1969).

3 The Tunneling Analogy

If the standard time-independent radial Schrödinger equation is written in units such that $\hbar^2/2\mu = 1$, where $h = 2\pi\hbar$ is Planck's constant and μ is the reduced mass, it then takes the form

$$-\frac{\mathrm{d}^2\psi(r)}{\mathrm{d}r^2} + [V(r) + \frac{l(l+1)}{r^2}]\psi(r) = E\psi(r), \tag{26}$$

where $V(r)$ is the potential energy function and E is the total energy. In particular, Eqs. (8) and (26) will be identical if we define the potential to be

$$V(r) = k^2[1 - n^2(r)], \tag{27}$$

and define $E = k^2$. The total (or effective) potential is the sum of the potential function $V(r)$ and the centrifugal potential, written as

$$V_l(r) = k^2[1 - n^2(r)] + \frac{l(l+1)}{r^2}. \tag{28}$$

(see Fig. 1).

Note that in distinction to the quantum mechanical situation, the effective potential is also a function of k. We consider (as an introduction to the methodology) the special case of a spherical particle with a constant refractive index n throughout (Johnson 1993). The potential in this case is given by

$$V_l(r) = \left\{ \begin{array}{cc} k^2(1-n^2) + l(l+1)/r^2, & r \le a \\ l(l+1)/r^2, & r > a \end{array} \right\} \tag{29}$$

where the values of n^2 and k^2 will define whether this potential is attractive or repulsive. We are interested in the case of a dielectric particle with $n > 1$.

For the specific example presented in Johnson (1993), we consider first the potential function $V_{40}(r)$ for a particle of radius a, the refractive index $n = 1.47$, and wavenumber $k = 33/a$. For convenience, the unit of length and the particle radius are chosen to be equal. Therefore, $a = 1$ and $k = 33$. As another example, we apply this technique to the case that $n = 4/3$ and $k = 33$. The effective potential for the partial wave $l = 40$ is designated $V_{40}(r)$, and details of these cases are shown in Fig. 2a–c (for $n = 1.47, l = 40$, with additional details later). The well in the region $r_1 < r < a$ for a given k is surrounded by the two classically 'forbidden' regions $0 \le r < r_1$ and $a < r < r_2$. The points r_1 and r_2 are called the classical turning

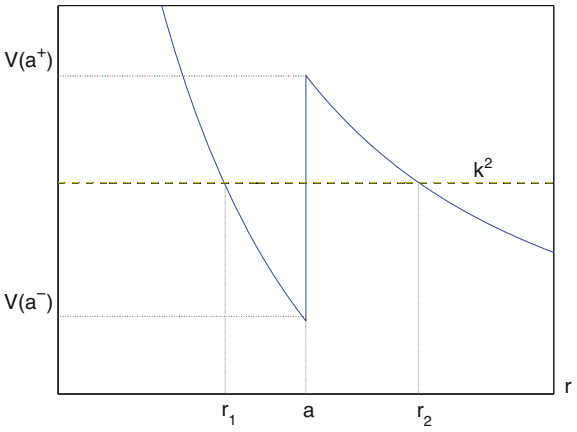


Fig. 1 Effective potential associated with a spherical dielectric particle

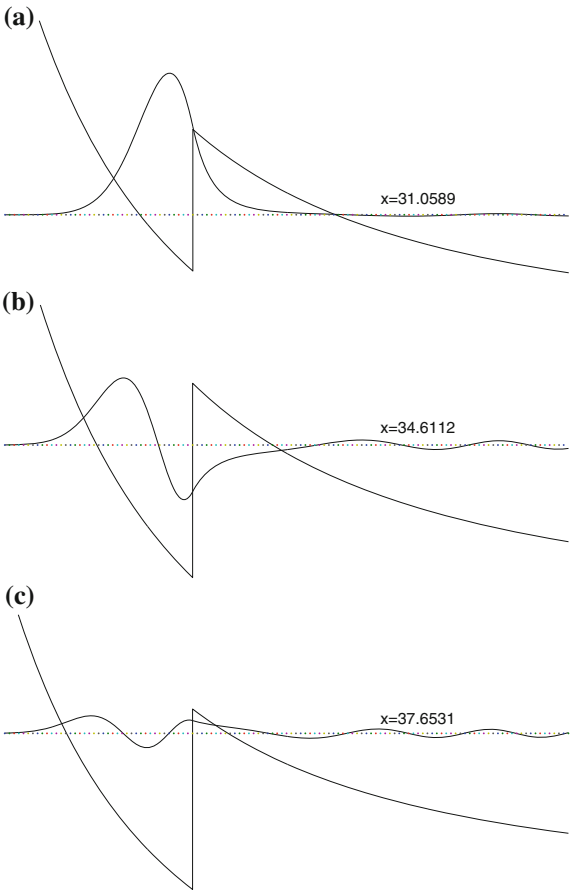


Fig. 2 Radial wave functions for three TE resonances (n = 1.47, l = 40)

points. They can be found by evaluating the local wavenumber $p_l(r)$, which is defined by $p_l^2(r) = E - V_l(r)$. This can be written in the form

$$p_l^2(r) = k^2 n^2(r) - \frac{l(l+1)}{r^2}. \tag{30}$$

In the corresponding quantum-mechanical problem, the particle can tunnel through the classically forbidden region $a < r < r_2$ into the classically allowed potential well. For specific values of the energy level, the wave will become temporally trapped in the well, oscillating back and forth many times, with part of the wave tunneling back each time it encounters the discontinuity at $r = a$. These phenomena are often referred to as quasi-bound states or resonances. The resonance described here can be called a shape resonance because the resonance behavior arises from the shape of the potential, i.e., the well and the barrier. Such resonances correspond to the usual interpretation of morphology-dependent resonances

(MDRs) in terms of electromagnetic 'rays' propagating around the inside surface of the sphere. Electromagnetic energy can tunnel through the classically forbidden region and become temporarily trapped in resonance states. In the following discussion, we assume that the refractive index is a real quantity (though this is not strictly necessary).

In Fig. 1, the shape of the potential well depends on the energy k^2 as noted previously. However, in the quantum-mechanical problem the potential is independent of the energy; this is clearly not the case here. Figure 1 shows the case when the energy k^2 lies between the top, $V(a^+)$, and bottom, $V(a^-)$, of the well. As noted in Johnson (1993), when k increases, the bottom of the potential well will drop. The energy level k^2 will finally coincide with the top of the well. In quantum mechanics, only certain levels of energy will satisfy the boundary conditions and are allowed in a potential well. The problem of shape resonances is similar in this regard. The boundary conditions at $r = 0$ are given by $S_l(0) = T_l(0) = 0$. These conditions are necessary to make sure that all scattering solutions are finite at the origin. The solution for the external region $r > a$ is given in terms of a linear combination of the Riccati–Bessel functions $\psi_l(kr)$ and $\chi_l(kr)$. In the classically forbidden region, $a < r < r_2$, these two functions have opposite behaviors. When the function $\psi_l(kr)$ is monotonically increasing rapidly in this region, the function $\chi_l(kr)$ is monotonically decreasing. At $r = r_2$, these functions stop their monotonic behavior and begin an oscillatory behavior in the outside region $r > r_2$.

To find a resonance, we impose the condition that the wave function monotonically decays in the barrier region, so it will tend to zero as the barrier is extended to $r \to \infty$. The quasi-bound state will then become a real bound state. Therefore, only the decreasing function $\chi_l(kr)$ is permitted in the barrier region. This implies that the coefficient that multiplies the increasing function $\psi_l(kr)$ in the wave function defined by (13) and (14) must be zero; i.e., $\beta_l = 0$ ($\alpha_l = 0$) at the location of a TE (TM) resonance, respectively. These conditions that are used to determine a shape resonance and to find the location of a MDR are the same. By substituting $\beta_l = 0$ and $\alpha_l = 0$ into (22) and (23) gives the equations that can be used to determine the locations of TE and TM resonances, respectively (though we only present the TE resonances in this chapter):

$$\psi_l(nx)\chi_l'(x) = n\psi_l'(nx)\chi_l(x), \tag{31}$$

$$\psi_l'(nx)\chi_l(x) = n\psi_l(nx)\chi_l'(x). \tag{32}$$

These equations have infinitely many discrete values of the size parameter $x_0 = ka$. However, only the finite number of values of x_0 that are in the range between the bottom and top of the potential well are considered to be resonant states. There are no solutions below the bottom of the well. The solutions above the top of the well are not classified as resonances, because they are too wide to have the properties discussed above. In Fig. 2a, the bottom and the top of the potential well (for the case considered by Johnson (1993), namely $n = 1.47$ and $l = 40$) are 27.5 and 40.5, respectively. This potential has three TE resonances between 27.5 and 40.5. The TE

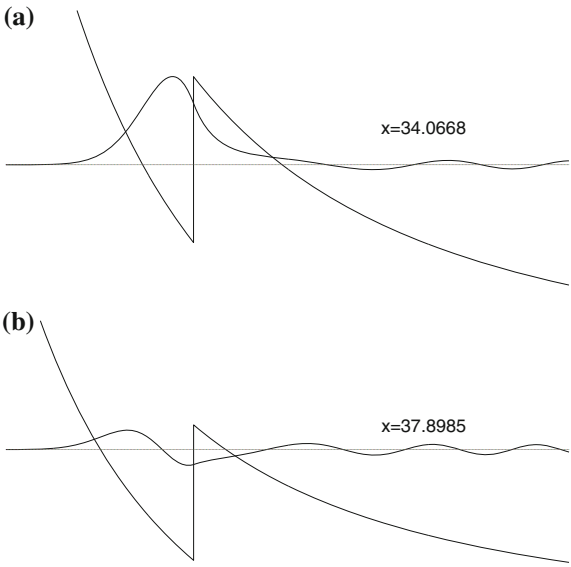


Fig. 3 Radial wave functions for two TE resonances (n = 4/3, l = 40)

resonances are located at 31.0589, 34.6112, and 37.6531. For the case $n = 4/3$ and $l = 40$ (Fig. 3a, b) the bottom and the top of the potential well are 30.365 and 40.5, respectively. This potential has two TE resonances, located at 34.0668 and 37.8985. In these Figures the wave functions are the Debye potential functions $S_{40}(r)$, obtained by solving Eq. (8). At the proper level, they are shown superimposed upon the potential function $V_{40}(r)$. (Note that solving (9) for $T_{40}(r)$ would give the TM resonances.) The wave functions show bound state within the region of potential well. The lowest-level wave function has a single peak, the next level has two peaks (positive and negative), and the third level has three peaks. Electromagnetic energy is temporarily trapped in the potential well. It can enter and leave the potential well by tunneling through the outer centrifugal barrier. The deeper well has a larger barrier than the upper levels. In Fig. 3a, b the wave functions for the two TE resonances for the case $n = 4/3$ and $l = 40$ are shown. They are the Debye potential functions $S_{40}(r)$ again obtained by solving Eq. (8). At the same proper level in the case $n = 1.47$ and $l = 40$, they are shown superimposed upon the potential function $V_{40}(r)$. The wave functions have only two resonances inside the region of potential well for this case. The lowest-level wave function has a single peak, and the next level has two peaks (positive and negative).

Figure 4a–c shows the change in resonance pattern that the wave function exhibits for the TE $l = 40$ mode located at $x_0 = 34.6112$. The top picture shows the wave function for the case $x > 34.6112$, which is above the resonance. The wave function $S_{40}(r)$ shows a monotone increase in the tunneling region. The amplitude of the wave function outside the particle $r > r_2$ is much larger than the amplitude inside the particle. The center picture shows the case $x = 34.6112$, for which the wave function decreases monotonically in the tunneling region. The amplitude of

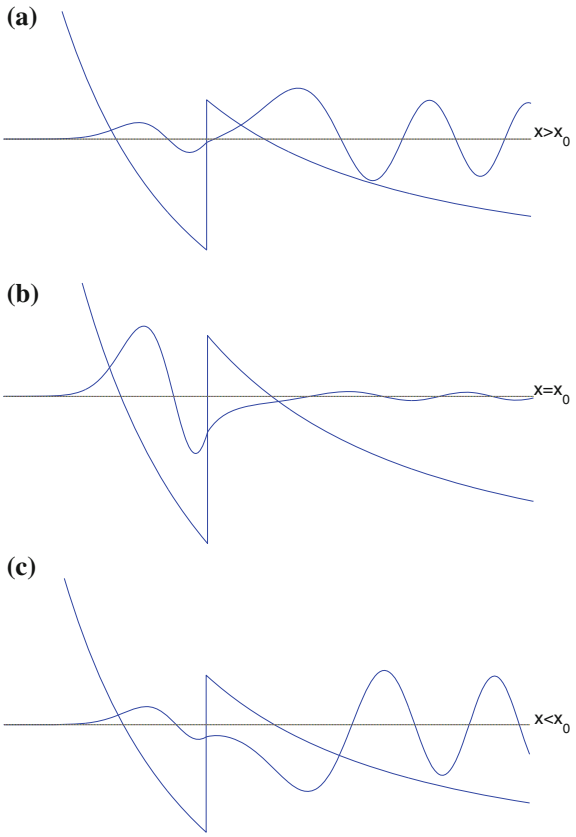


Fig. 4 Behavior of the TE wave function in the vicinity of a resonance: behavior for a size parameter value slightly above resonance (**a**); on resonance (**b**); below resonance (**c**)

the wave function inside the particle is much larger than the amplitude outside because the field strength increases rapidly in the layer just outside the surface, $a < r < r_2$, and then continues to a maximum inside the particle near the surface (both inside and outside). This is sufficient to define a resonance. The bottom picture shows the case $x < 34.6112$, which is below the resonance. The wave function increases monotonically in the tunneling region. Thus, the amplitude inside a particle is smaller than the amplitude outside. This case is similar to the case $x > 34.6112$ (top picture) except that the wave amplitude decreases in this region before increasing.

4 The Two-Layer Inhomogeneous Sphere

We return to the radial Debye potentials $S_l(r)$ and $T_l(r)$ [which satisfy respectively the second-order differential Eqs. (8) and (9)]. In this case, we consider a piecewise constant refractive index associated with a two-layer dielectric sphere. In regions

where the indices of refraction have the constant value n_1 and n_2, the two radial Schrödinger equations have the same form and the linearly independent solutions are again the Riccati–Bessel functions

$$\psi_l(nkr) = nkrj_l(nkr), \tag{33}$$

$$\chi_l(nkr) = nkry_l(nkr). \tag{34}$$

Here $j_l(nkr)$ and $y_l(nkr)$ are spherical Bessel functions of the first and second kinds, respectively. Since we need the conditions defining the locations of the resonances of TE and TM modes in this two-layer case, we again examine the solutions of Eqs. (8) and (9). Since the sphere is embedded in an external medium, we have three piecewise-uniform regions to consider.

4.1 TE Mode

For the TE mode of this model the solutions are as follows:

$$\text{Region 1}(0 < r < R_1)\text{: } S_{1l}(r) = \psi_l(n_1kr), \tag{35}$$

$$\text{Region 2 } (R_1 < r < R_2)\text{: } S_{2l}(r) = A_l[\chi_l(n_2kr) + \alpha_l\psi_l(n_2kr)], \tag{36}$$

$$\text{Region 3 } (r > R_2)\text{: } S_{3l}(r) = B_l[\chi_l(kr) + \beta_l\psi_l(kr)]. \tag{37}$$

Matching solutions at $r = R_1$ and $r = R_2$ using the log-derivative formalism in Eqs. (15) and (16) for the continuity of the solution, where the refractive index is discontinuous we obtain

$$A_l = \frac{\psi_l(n_1kR_1)}{\chi_l(n_2kR_1) + \alpha_l(n_2kR_1)}, \tag{38}$$

$$B_l = \frac{A_l[\chi_l(n_2kR_2) + \alpha_l\psi_l(n_2kR_2)]}{\chi_l(kR_2) + \beta_l\psi_l(kR_2)}, \tag{39}$$

where

$$\alpha_l = -\frac{n_2\psi_l(n_1kR_1)\chi_l'(n_2kR_1) - n_1\psi_l'(n_1kR_1)\chi_l(n_2kR_1)}{n_2\psi_l(n_1kR_1)\psi_l'(n_2kR_1) - n_1\psi_l'(n_1kR_1)\psi_l(n_2kR_1)}, \tag{40}$$

$$\beta_l = -\frac{\chi_l'(kR_2)[\chi_l(n_2kR_2) + \alpha_l\psi_l(n_2kR_2)] - n_2\chi_l(kR_2)[\chi_l'(n_2kR_2) + \alpha_l\psi_l'(n_2kR_2)]}{\psi_l'(kR_2)[\chi_l(n_2kR_2) + \alpha_l\psi_l(n_2kR_2)] - n_2\psi_l(kR_2)[\chi_l'(n_2kR_2) + \alpha_l\psi_l'(n_2kR_2)]}. \tag{41}$$

4.2 TM Mode

For completeness, we state the corresponding solutions for the TM mode.

$$\text{Region 1 } (0<r<R_1):\ T_{1l}(r)=\psi_l(n_1kr), \tag{42}$$

$$\text{Region 2 } (R_1<r<R_2):\ T_{2l}(r)=\bar{A}_l[\chi_l(n_2kr)+\bar{\alpha}_l\psi_l(n_2kr)], \tag{43}$$

$$\text{Region 3 } (r>R_2):\ T_{3l}(r)=\bar{B}_l[\chi_l(kr)+\bar{\beta}_l\psi_l(kr)]. \tag{44}$$

Again matching the solutions in each region at the boundaries $r=R_1$ and $r=R_2$, we obtain

$$\bar{A}_l=\frac{\psi_l(n_1kR_1)}{\chi_l(n_2kR_1)+\bar{\alpha}_l(n_2kR_1)}, \tag{45}$$

$$\bar{B}_l=\frac{\bar{A}_l[\chi_l(n_2kR_2)+\bar{\alpha}_l\psi_l(n_2kR_2)]}{\chi_l(kR_2)+\bar{\beta}_l\psi_l(kR_2)}, \tag{46}$$

where

$$\bar{\alpha}_l=-\frac{n_2\psi_l'(n_1kR_1)\chi_l(n_2kR_1)-n_1\psi_l(n_1kR_1)\chi_l'(n_2kR_1)}{n_2\psi_l'(n_1kR_1)\psi_l(n_2kR_1)-n_1\psi_l(n_1kR_1)\psi_l'(n_2kR_1)}, \tag{47}$$

$$\bar{\beta}_l=-\frac{n_2\chi_l'(kR_2)[\chi_l(n_2kR_2)+\bar{\alpha}_l\psi_l(n_2kR_2)]-\chi_l(kR_2)[\chi_l'(n_2kR_2)+\bar{\alpha}_l\psi_l'(n_2kR_2)]}{n_2\psi_l'(kR_2)[\chi_l(n_2kR_2)+\bar{\alpha}_l\psi_l(n_2kR_2)]-\psi_l(kR_2)[\chi_l'(n_2kR_2)+\bar{\alpha}_l\psi_l'(n_2kR_2)]}. \tag{48}$$

4.3 The Potential Function

The potential $V_l(r)$ is now defined as

$$V_l(r)=\left\{\begin{array}{ll} k^2(1-n_1^2)+l(l+1)/r^2, & 0\le r\le R_1,\\ k^2(1-n_2^2)+l(l+1)/r^2, & R_1\le r\le R_2,\\ l(l+1)/r^2, & r>R_2. \end{array}\right\} \tag{49}$$

This corresponds to the refractive index profile:

$$n(r)=\left\{\begin{array}{ll} n_1, & 0\le r\le R_1,\\ n_2, & R_1\le r\le R_2,\\ 1, & r>R_2. \end{array}\right. \tag{50}$$

The relative values of n_1 and n_2 define the nature of the potential, i.e., whether it is attractive or repulsive. We focus here on a dielectric particle with $n_1 > 1$, $n_2 > 1$. Using a similar technique to that in Sect. 2, we consider the potential function $V_l(r)$ for the two-layer sphere with radii R_1 and R_2 (and corresponding refractive indices n_1 and n_2). Now the size parameters are $x = kR_1$ and $y = kR_2$, respectively (note that y as defined here should not be confused with that is Eq. (12) defining the second Riccati–Bessel function). For convenience, the unit of length and the particle radius are again chosen to be equal. We now separate the refractive index profile into two obvious and distinct cases.

4.3.1 Case 1 and Case 2

Case 1: Increasing refractive index profile

$$n(r) = \begin{cases} n_1, & 0 \le r \le R_1, \\ n_2(>n_1), & R_1 \le r \le R_2, \\ 1, & r > R_2. \end{cases} \tag{51}$$

Case 2: Decreasing refractive index profile

$$n(r) = \begin{cases} n_1, & 0 \le r \le R_1, \\ n_2(<n_1), & R_1 \le r \le R_2, \\ 1, & r > R_2. \end{cases} \tag{52}$$

Figures 7 and 8 show the generic form of potential function $V_{40}(r)$ given by Eq. (49) corresponding to the profiles (51) and (52). Figures 9a–c and 10a–b show the wave function $S_{40}(r)$ for a particle with radii $R_1 = x_0/k$ and $R_2 = 1$, and refractive indices $n_1 = 1.2$ and $n_2 = 1.5$ (representing an increasing refractive index profile as shown in Fig. 5) with wavenumbers $k = x_0/R_1 = y_0/R_2$. Figures 11a, b and 12 show similar results for a particle of radius $R_1 = x_0/k$ and $R_2 = 1$, refractive index $n_1 = 1.52$, and $n_2 = 1.25$ (representing a decreasing refractive index profile as shown in Fig. 6) with wavenumbers $k = x_0/R_1 = y_0/R_2$.

For specific values of the energy level, as noted earlier, the wave will become temporally trapped in the well, oscillating back and forth many times, with part of the wave tunneling back each time it encounters the discontinuity at $r = a$. This is where the shape resonances of the two-layer model were found. Note that in all cases the potential well depends on the energy k^2 and the refractive index in each region. The figures show the case when the energy k^2 lies between the top and the bottom of the well. Obviously, the shape of the well depends on the values of

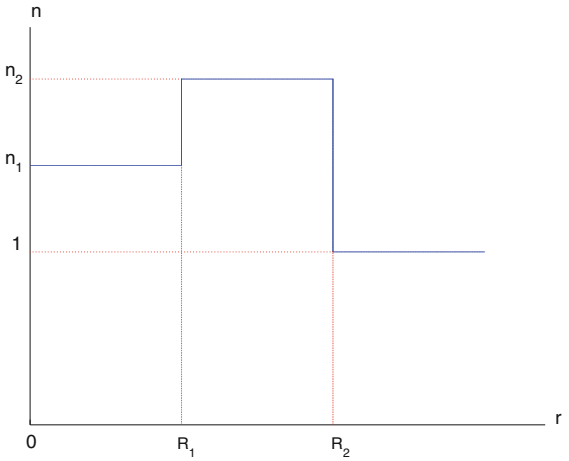


Fig. 5 Case 1: Refractive index profile for $n_1 < n_2$

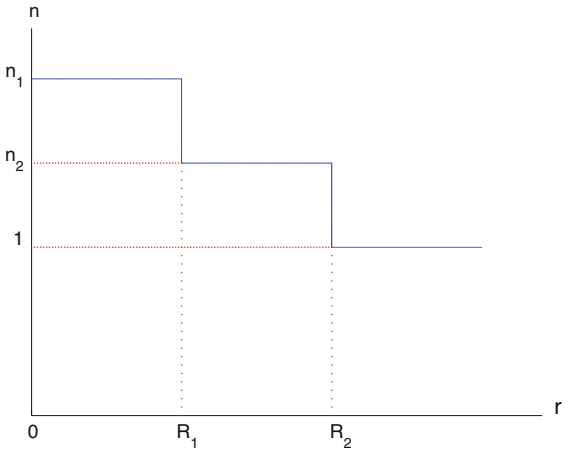


Fig. 6 Case 2: Refractive index profile for $n_1 > n_2$

refractive indices n_1 and n_2 in each region. When $n_1 < n_2$ (as in Fig. 5), we have a single well (see Fig. 7); however, when $n_1 > n_2$ we have double wells (as in Fig. 8). Similar to the case of the uniform model, the boundary conditions at $r = 0$ are given by $S_{1l}(0) = T_{1l}(0) = 0$ to guarantee that all scattering solutions are finite at the origin. The solutions are linear combinations of the Riccati–Bessel functions $\psi_l(n_2 kr)$ and $\chi_l(n_2 kr)$ in Region 2 and $\psi_l(kr)$ and $\chi_l(kr)$ in Region 3. For the case that $n_1 < n_2$, when n_2 is much larger than n_1, the well will be deeper and wider. The most striking feature of this case with the potential function $V_{40}(r)$ (for increasing refractive index) is the presence of a potential well in the region $a_1 < r < R_2$. This is

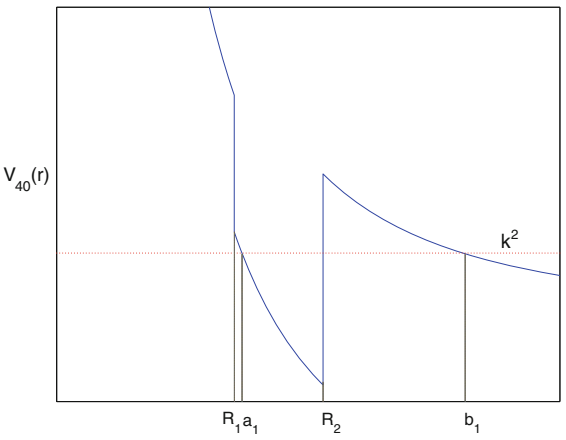


Fig. 7 Case 1: $V(r)$ potential for $n_1 < n_2$

a classically allowed region. The well is surrounded by the two classically forbidden regions $0 \leq r < a_1$ and $R_2 < r < b_1$. The points a_1 and b_1 are called the classical turning points. In the equivalent quantum-mechanical problem a particle can tunnel through the classically forbidden region $R_2 < r < b_1$, into the classically allowed potential well. In the case that $n_1 > n_2$, the most interesting feature of this case is the presence of the potential wells in the regions $a_2 < r < R_1$ and $b_2 < r < R_2$. These are the classically allowed regions. They are surrounded by the three classically forbidden regions $0 \leq r < a_2$, $R_1 < r < b_2$, and $R_2 < r < c_2$. The points a_2, b_2, and c_2 are the classical turning points for this case. In the forbidden regions $r < b_1$ in Fig. 7 and $r < c_2$ in Fig. 8, the two functions $\psi_l(kr)$ and $\chi_l(kr)$ have opposite behaviors. When the function $\psi_l(kr)$ increases rapidly (again monotonically) in this region, the function $\chi_l(kr)$ has monotone decreasing behavior. The classical turning points for the exhibited value of k^2 (dotted line) are denoted by a_1, b_1 and a_2, b_2, c_2, respectively, in these figures. Of course, both R_1 and R_2 are classical turning points also.

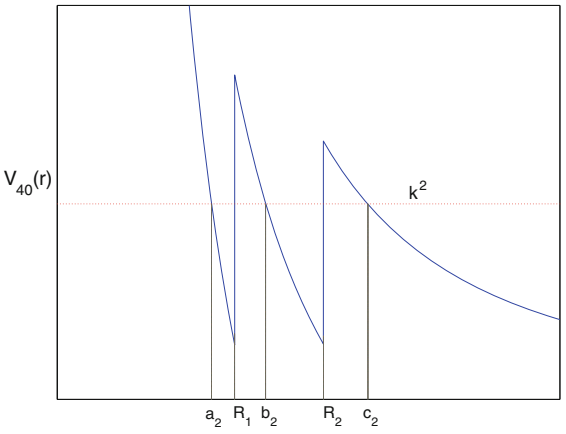


Fig. 8 Case 2: $V(r)$ potential for $n_1 > n_2$

To find the resonance, as before we use the (sufficient) condition that the wave function decays monotonically in the barrier region to ensure that it goes to zero as the barrier radius $r \to \infty$. Therefore, we need only the monotonically decreasing function $\chi_l(kr)$ in the barrier region. By substituting $\alpha_l = 0$ and $\beta_l = 0$, $\bar{\alpha}_l = 0$ and $\bar{\beta}_l = 0$ into Eqs. (40), (41), (47), and (48), we have the equations determining locations of TE and TM resonances, respectively:

TE:

$$n_2\psi_l(n_1kR_1)\chi_l'(n_2kR_1) = n_1\psi_l'(n_1kR_1)\chi_l(n_2kR_1), \tag{53}$$

$$\chi_l'(kR_2)\chi_l(n_2kR_2) = n_2\chi_l(kR_2)\chi_l'(n_2kR_2). \tag{54}$$

TM:

$$n_2\psi_l'(n_1kR_1)\chi_l(n_2kR_1) = n_1\psi_l(n_1kR_1)\chi_l'(n_2kR_1), \tag{55}$$

$$n_2\chi_l'(kR_2)\chi_l(n_2kR_2) = \chi_l(kR_2)\chi_l'(n_2kR_2). \tag{56}$$

These equations have infinitely many discrete values of the size parameters x_0 and y_0. However, only the finite number of values of x_0 and y_0 that are in the range between the top and the bottom of the potential well are considered to be resonant states. Similar to the two-layer model, there are no solutions below the bottom of the well and above the top of the well. Figure 9a–c shows the potential function $V_{40}(r)$ and the wave function $S_{40}(r)$ for $n_1 = 1.2$, $n_2 = 1.5$, and $l = 40$. This potential supports three TE resonances for specific value of x_0. They are located at $x_0 = 30.3828$ with $y_0 = 32.2993$ (Fig. 9a); $y_0 = 35.4868$ (Fig. 9b) and $y_0 = 38.3548$ (Fig. 9c). Figure 10a, b shows the potential function $V_{40}(r)$ and the wave function $S_{40}(r)$ for $n_1 = 1.2$, $n_2 = 1.5$, and $l = 40$. The TE resonances for this example are located at $x_0 = 34.7905$ with $y_0 = 35.4868$ (Fig. 10a), and $y_0 = 38.3548$ (Fig. 10b). Figure 11a, b shows the potential function $V_{40}(r)$ and the wave function $S_{40}(r)$ for $n_1 = 1.52$, $n_2 = 1.25$, and $l = 40$. The two TE resonances are located at $x_0 = 29.5815$ with $y_0 = 33.5645$ (Fig. 11a), and $y_0 = 38.2446$ (Fig. 11b). Figure 12 shows the potential function $V_{40}(r)$ and the wave function $S_{40}(r)$ for $n_1 = 1.52$, $n_2 = 1.25$, and $l = 40$. The location for the TE resonance is at $x_0 = 33.4975$ with $y_0 = 38.2446$. Figures 13a–c and 14a–c show the change in form that the wave function experiences as the system transverses the TE for both cases of $n_1 < n_2$ and $n_1 > n_2$ with $l = 40$, respectively. In Fig. 13b, the TE resonance is located at $x_0 = 30.3828$ and $y_0 = 32.2993$. Figure 13a shows the case for $x = x_0$ and $y = 32.0993 < y_0$ which is below the resonance. Figure 13c shows the wave function for the case $x = x_0$ and $y = 32.4993 > y_0$ which is above the resonance. In Fig. 14b, the TE resonance is located at $x_0 = 29.5815$ and $y_0 = 33.5645$. Figure 14a shows the case for $x = x_0$ and $y = 33.0645 < y_0$, which is below the resonance. Figure 14c shows the wave function for the case $x = x_0$ and $y = 33.9645 > y_0$, which is above the resonance.

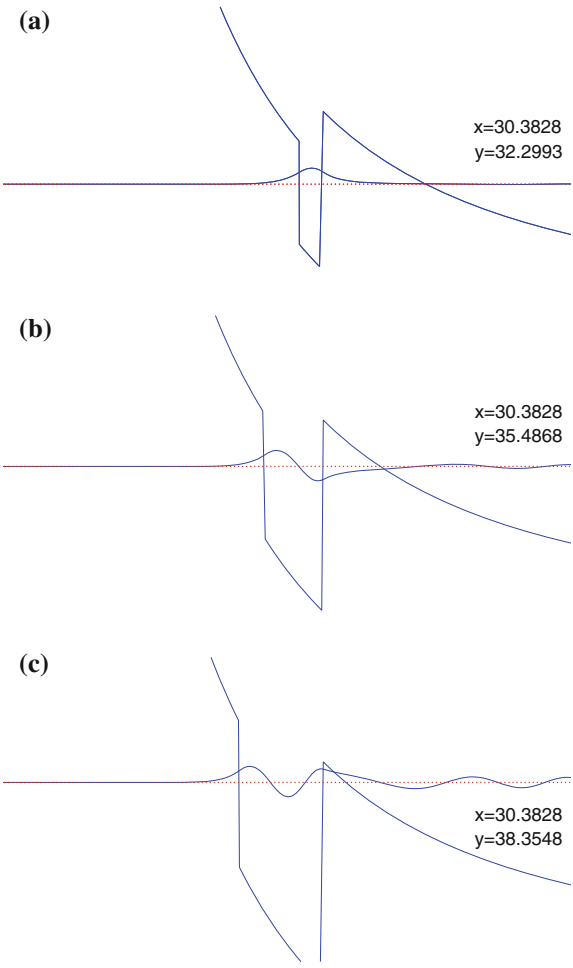


Fig. 9 Radial wave functions for three TE resonances with $x = 30.3828$; $y = 32.2993$ (**a**); $x = 30.3828$; $y = 35.4868$ (**b**); $x = 30.3828$; $y = 38.3548$ (**c**), corresponding to the refractive index profile in Case 1 ($n_1 = 1.2$, $n_2 = 1.5$, $l = 40$)

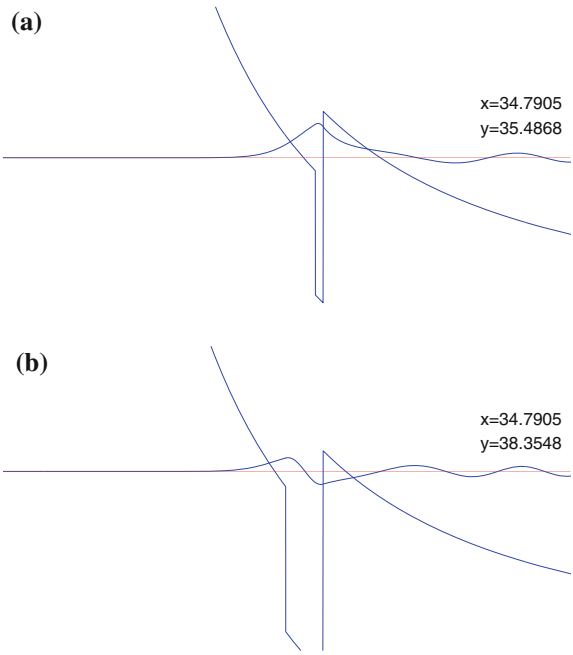


Fig. 10 Radial wave functions for two TE resonances with $x = 34.7905$; $y = 35.4868$ (**a**); $x = 34.7905$; $y = 38.3548$ (**b**), corresponding to the refractive index profile in Case 1 ($n_1 = 1.2$, $n_2 = 1.5$, $l = 40$)

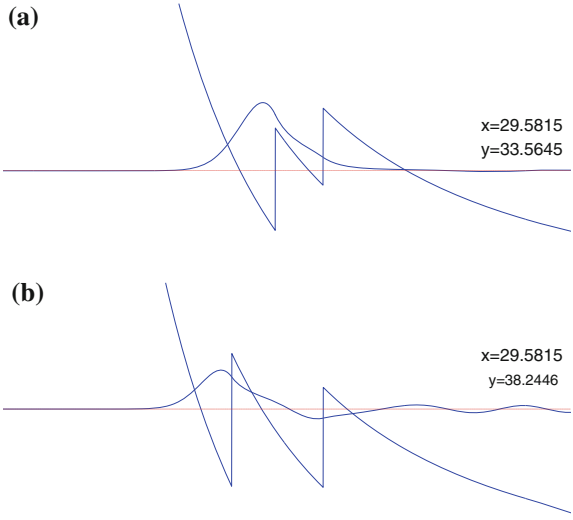


Fig. 11 Radial wave functions for two TE resonances with $x = 29.5815$; $y = 33.5645$ (**a**); $x = 29.5815$; $y = 38.2446$ (**b**), corresponding to the refractive index profile in Case 2 ($n_1 = 1.52$, $n_2 = 1.25, l = 40$)

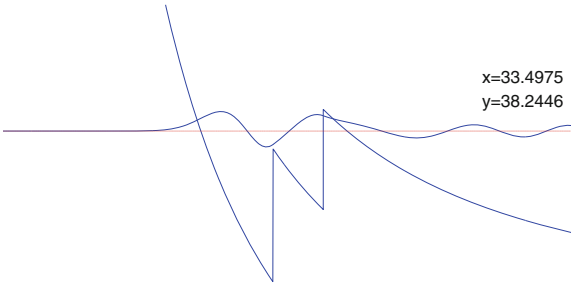


Fig. 12 Radial wave function for a TE resonance with $x = 33.4975$; $y = 38.2446$, corresponding to the refractive index profile in Case 2 ($n_1 = 1.52$, $n_2 = 1.25, l = 40$)

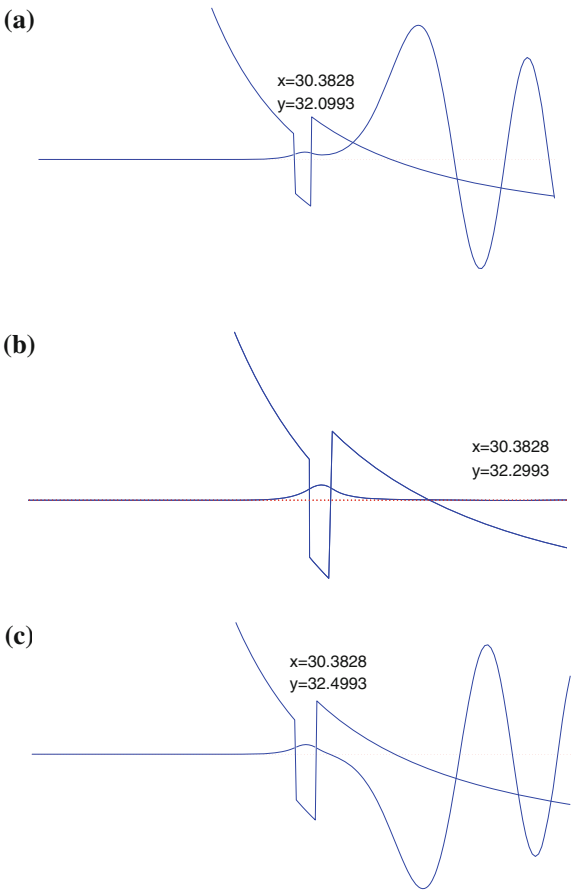


Fig. 13 Behavior of the TE wave function in the vicinity of a resonance for the case $n_1 < n_2$; for $x = 30.3828$, the behavior for a size parameter value slightly below resonance **a** $y = 32.0993$; on resonance **b** $y = 32.2993$; above resonance **c** $y = 32.4993$

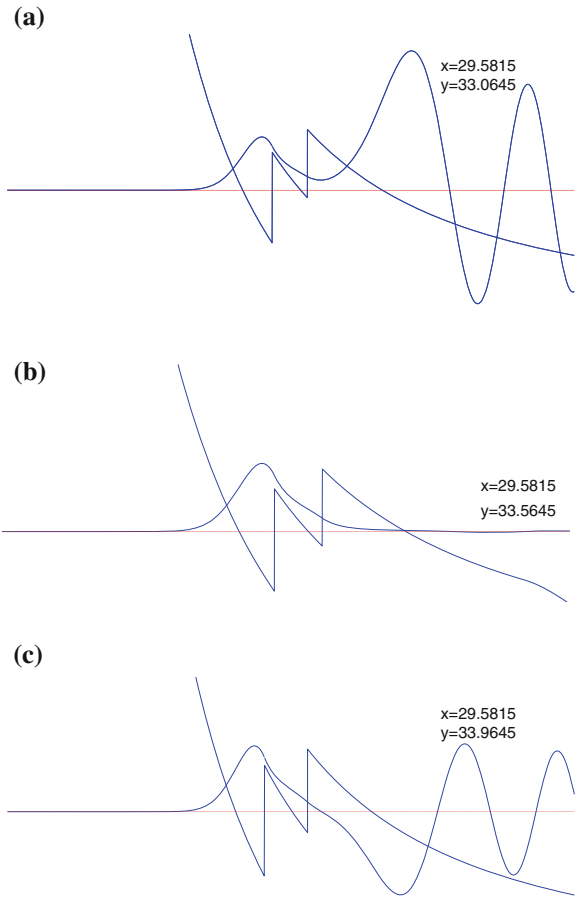


Fig. 14 Behavior of the TE wave function in the vicinity of a resonance for the case $n_1 > n_2$; for $x = 29.5815$, the behavior for a size parameter value slightly below resonance **a** $y = 33.0645$; on resonance **b** $y = 33.5645$; above resonance **c** $y = 33.9645$

5 Conclusion

In this chapter, we have examined the interaction of plane electromagnetic waves in piecewise continuous radially inhomogeneous media. We have summarized a major part of the seminal work of Johnson (1993), who discussed in detail so-called MDRs for a uniform sphere embedded in a different uniform medium. In a strict sense, his work can be interpreted as a study of plane wave electromagnetic scattering in piecewise constant media in $[0, \infty)$. Complementary material (also based on Johnson's 1993 paper) was presented in Sect. 3, where the relationship between the radial electromagnetic wave equations and the corresponding time-independent radial Schrödinger equation is made. Essentially the former problem can be regarded as the latter with a wavenumber-dependent potential. In Sect. 4, the TE

analysis of Johnson was extended to piecewise constant two-layer spheres (with refractive indices n_1 and n_2) in a basic approach to modeling inhomogeneity, for both $n_1 > n_2$ and $n_1 < n_2$. The corresponding theory for the TM modes was also developed.

Acknowledgments The authors would like to thank the reviewer for his very helpful and constructive comments that resulted in a much improved version of the original manuscript.

References

Adam JA (2013) "Rainbows" in homogeneous and radially inhomogeneous spheres: connections with ray, wave and potential scattering theory, Chap. 3. In: Bourama T (ed) Advances in interdisciplinary mathematical research: applications to engineering, physical and life sciences, vol 37. Springer, Berlin

Adam JA (2014) Scattering of electromagnetic plane waves in radially inhomogeneous media: ray theory, exact solutions and connections with potential scattering theory, Chap. 3. In: Kokhanovsky A (ed) Light scattering reviews, vol 9. Springer, Berlin

Adams MJ (1981) An introduction to optical waveguides. Wiley, New York

Alexopoulos NG (1971) Scattering from inhomogeneous lenses with singular points. J Opt Soc Am 61:876–878

Alexopoulos NG (1972) Scattering from inhomogeneous cylindrically symmetric lenses with a line infinity in the index of refraction. J Opt Soc Am 62:1088–1094

Alexopoulos NG (1974) On the refractive properties of media with poles or zeros in the index of refraction. IEEE Trans Antennas Propag AP-22:242–251

Bohren CF, Huffman DR (1983) Absorption and scattering of light by small particles. Interscience, New York

Brockman CL (1974) High frequency electromagnetic wave backscattering from radially inhomogeneous dielectric spheres. PhD dissertation, University of California, Los Angeles

Brunsting A, Mullaney PF (1972) Light scattering from coated spheres: model for biological cells. Appl Opt 11:675–680

Cornbleet S (1984) Microwave and optical ray geometry. Wiley, New York

Fenn RW, Oser H (1965) Scattering properties of concentric soot-water spheres for visible and infrared light. Appl Opt 4:1504–1509

Hill SC, Benner RE (1988) Morphology-dependent resonances. In: Barber PW, Chang, RK (eds) Optical effects associated with small particles. World Scientific, Singapore

Johnson BR (1993) Theory of morphology-dependent resonances: shape resonances and width formulas. J Opt Soc Am A10:343–352

Kerker M (1969) The scattering of light and other electromagnetic radiation. Academic, New York

Kitchen JC, Zaneveld JRV (1992) A three-layered sphere model of the optical properties of phytoplankton. Limnol Oceanogr 37(8):1680–1690

Leonhardt U (2006) Optical conformal mapping. Science 312:1777–1780

Leonhardt U, Philbin T (2010) Geometry and light: the science of invisibility. Dover Publications, New York

Nuntaplook U (2013) Topics in electromagnetic, acoustic and potential scattering theory. PhD dissertation, Old Dominion University

Nuntaplook U, Adam JA (2014) Scalar wave scattering by two-layer radial inhomogeneities. Appl Math E-Notes 14(2014):185–192

Pendry JB, Schurig D, Smith DR (2006) Controlling electromagnetic fields. Science 312:1780–1782

Pohrivchak MA (2014) Ray-and wave-theoretic approach to electromagnetic scattering from radially inhomogeneous spheres and cylinders. PhD Dissertation, Old Dominion University
van de Hulst HC (1981) Light scattering by small particles. Dover, New York
Wyatt PJ (1962) Scattering of electromagnetic plane waves from inhomogeneous spherically symmetric objects. Phys Rev 127:1837–1843

Light Scattering and Thermal Emission by Primitive Dust Particles in Planetary Systems

Hiroshi Kimura, Ludmilla Kolokolova, Aigen Li and Jérémy Lebreton

1 Introduction

Stardust grains are tiny solid samples of stars, newly condensed in an expanding atmosphere of a dying star and injected into interstellar space by stellar wind and radiation pressure. Such a dust grain in the interstellar medium goes back and forth between diffuse clouds and dense clouds before it experiences star formation at the core of a dense cloud (Greenberg 1984). The majority of prestellar interstellar grains evaporate during star formation, but circumstellar gas condenses as pristine dust grains when a gaseous envelope of a newly born star cools down to the melting temperatures of solids (Keller and Messenger 2011). In consequence, it is no wonder that interstellar grains are chemically similar to primitive dust particles in planetary systems, if the particles are least processed (Kimura 2013). Pristine dust grains coagulate together and form fluffy agglomerates of the grains in a proto-

H. Kimura (✉)
Graduate School of Science, Kobe University, C/O CPS (Center for Planetary Science), Chuo-Ku Minatojima Minamimachi, 7-1-48, Kobe 650-0047, Japan
e-mail: hiroshi_kimura@cps-jp.org

L. Kolokolova
Planetary Data System Group, Department of Astronomy, University of Maryland, Rm. 2337, Computer and Space Science Bldg, College Park, MD 20742, USA

A. Li
Department of Physics and Astronomy, University of Missouri, 314 Physics Building, Columbia, MO 65211, USA

J. Lebreton
Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA

J. Lebreton
NASA Exoplanet Science Institute, California Institute of Technology, 770 S. Wilson Ave, Pasadena, CA 91125, USA

© Springer-Verlag Berlin Heidelberg 2016

A. Kokhanovsky (ed.), *Light Scattering Reviews, Volume 11*,
Springer Praxis Books, DOI 10.1007/978-3-662-49538-4_8

planetary disk around a young star (Weidenschilling et al. 1989). In 1970s, a collection of interplanetary dust particles (IDPs), which originate from comets and/or asteroids, revealed that the particles are indeed fluffy agglomerates of submicron constituent grains (Brownlee et al. 1976). Collisional growth of fluffy agglomerates ends up with planetesimals, which are at present observed as comets and asteroids in planetary systems around main-sequence stars (Weidenschilling 1997). Therefore, planetesimals (i.e., comets and asteroids) are time capsules of primitive dust particles, which are reservoirs of information on the time of planetary system formation. At the present time, primitive dust particles released from comets and asteroids have been observed in the Solar System and debris disks[1] around other main-sequence stars through stellar radiation scattered or reradiated by the particles. The interpretation of observational data is, however, often not straightforward, unless light-scattering and thermal-emission properties of the particles are well known a priori.

Numerical simulation is a flexible and powerful approach to deducing light-scattering and thermal-emission properties of primitive dust particles in planetary systems from astronomical observations. In this review, we focus on numerical approaches to light scattering and thermal emission of fluffy agglomerates that simulate primitive dust particles in the Solar System and debris disks. In consequence, it must be emphasized that dust particles in protoplanetary disks around pre-main-sequence stars go beyond the scope of this review. This paper is organized as follows: we review and comment on popular models of dust agglomerates in Sect. 2, light-scattering techniques applicable to agglomerates in Sect. 3, light scattering by agglomerates in Sect. 4, thermal emission from agglomerates in Sect. 5, and integral optical quantities of agglomerates such as bolometric albedo, radiation pressure, and equilibrium temperature in Sect. 6. Furthermore, we pass our concluding remarks in Sect. 7 and finally provide a summary in Sect. 8.

2 Models of Dust Agglomerates in Planetary Systems

2.1 Artificial Configuration of Constituent Grains

There are a variety of ways to numerically configure constituent grains in an agglomerate, if the formation process of the agglomerate is not taken into account. The most simple algorithm for creating a fluffy agglomerate in control of its porosity would be to first form a cubic or a sphere with subvolumes and then

[1]A debris disk is defined as a dust disk that surrounds a main-sequence star with its age exceeding the lifetime of the disk. Therefore, the disk is not primordial, but must be replenished with dust particles released from their parent bodies such as asteroids and comets.

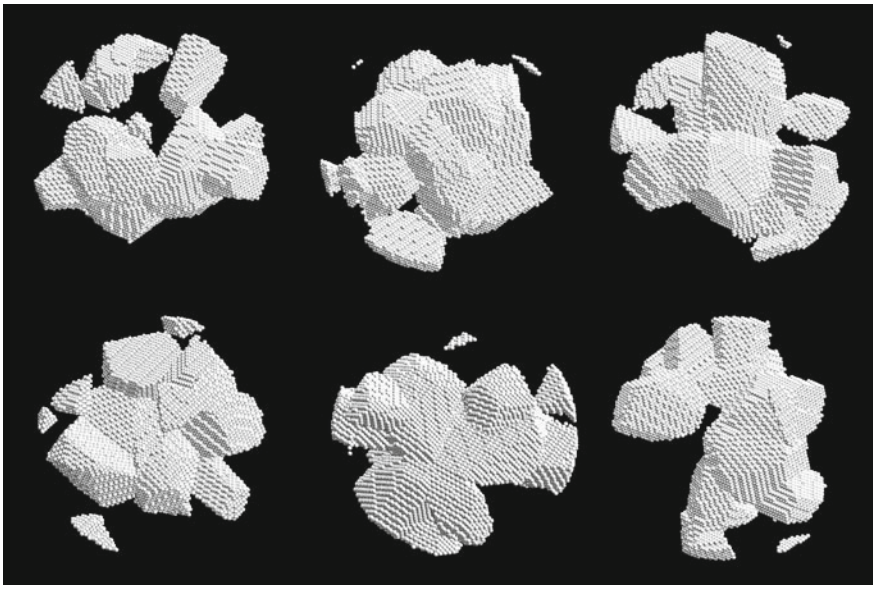


Fig. 1 Six realizations of dust agglomerates that formed by the removal of subvolumes consecutively from an initially spherical volume. The same algorithm was applied to generate the agglomerates, while different random numbers were used for their respective operations. From Zubko et al. (2009)

remove a portion of the subvolumes according to the desired porosity (e.g., Hage and Greenberg 1990; Zubko et al. 2009; Petrova and Tishkovets 2011; Kirchschlager and Wolf 2013). This algorithm has the advantage of easily producing irregularly shaped constituent grains, but the disadvantage is in producing equi-dimensional agglomerates. Moreover, there is a risk that an algorithm of this kind separates some of the constituent grains from the agglomerate, even in the case that the porosity of the agglomerate is not extremely high (see Fig. 1).

Alternatively, one could first create constituent grains in preferable sizes and shapes and then assemble a cluster of them in an arbitrary configuration. In this way, numerical calculations of electromagnetic waves scattered not only by a cluster of spheres, but also by a cluster of tetrahedrons, have been implemented to simulate light scattering by dust agglomerates in planetary systems (Xing and Hanner 1997; Yanamandra-Fisher and Hanner 1999). It is certainly of great interest from a mathematical point of view to consider arbitrary configuration and shape of constituent grains in dust agglomerates. However, one should keep in mind that the application of such an agglomerate to dust particles in planetary systems is not necessarily justified.

2.2 Coagulation Growth of Pristine Constituent Grains

Agglomerates of submicron constituent grains make up the anhydrous chondritic porous (CP) subset of IDPs collected in the stratosphere as well as Antarctic ice and snow (Brownlee 1985; Noguchi et al. 2015). CP IDPs are found to consist of chondritic grains embedded in organic-rich carbonaceous material and most likely originate from short-period comets. As a result, their chemical composition bears strong resemblance to the composition of cometary dust measured in situ for comet 1P/Halley as shown in Fig. 2 (Jessberger 1999; Kimura et al. 2003a). A dust agglomerate consisting of submicron grains is a natural consequence of dust growth due to collisions of submicron pristine condensates in a protoplanetary disk. If the collision takes place at a low relative velocity ($v \leq 1\,\mathrm{ms}^{-1}$), an agglomerate grows without restructuring of constituent grains and possess a fractal geometry (Wurm and Blum 1998). The structure of such an agglomerate is characterized by a fractal dimension D, which is defined as

$$m \propto a_{\mathrm{g}}^{D}, \tag{1}$$

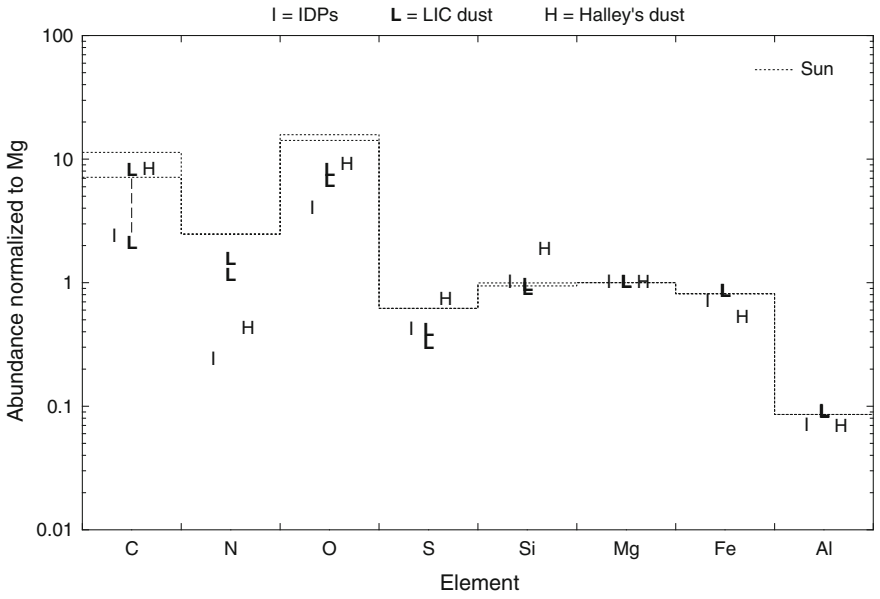


Fig. 2 The chemical compositions of cometary dust, interplanetary dust, interstellar dust as well as the solar photosphere, normalized to their Mg abundances. The data for cometary dust (H), interplanetary dust (I), and interstellar dust (L) correspond to those of dust in comet 1P/Halley, the chondritic porous subset of interplanetary dust particles, and dust in the Local Interstellar Cloud, respectively. The *two dotted lines* for the photosphere of the Sun reflect uncertainties in the solar photospheric composition, which also affect the composition of interstellar dust. From Kimura et al. (2003a)

where m and a_g are the mass and gyration radius of the agglomerate, respectively. The radius of gyration a_g is given by

$$a_g = \left[\frac{1}{2N^2} \sum_i^N \sum_j^N (\mathbf{r}_i - \mathbf{r}_j)^2 \right]^{1/2}, \tag{2}$$

where $\mathbf{r}_i$ and $\mathbf{r}_j$ are the position vectors of the i-th and j-th constituent grains, respectively, and N is the total number of the grains. At a higher velocity, a collision results in restructuring constituent grains and highly compressed agglomerates, or even destructing agglomerates (Dominik and Tielens 1997). Even at the maximum compression of agglomerates due to high-velocity collisions, possession of a fractal structure has been manifested by highly sophisticated numerical simulations (Wada et al. 2008). Therefore, it is reasonable as a first step to assume fractal agglomerates consisting of submicron grains when simulating light scattering and thermal emission of primitive dust particles in planetary systems.

If constituent grains in an agglomerate are identical, the number N of the grains and the characteristic radius a_c of the agglomerate fulfill the following relation:

$$N = k_0 \left(\frac{a_c}{a_0} \right)^D, \tag{3}$$

where a_0 is the radius of the constituent grains and k_0 is a proportionality constant of order unity, which depends on the coagulation process (see Fig. 3). The characteristic radius a_c of an agglomerate is defined by

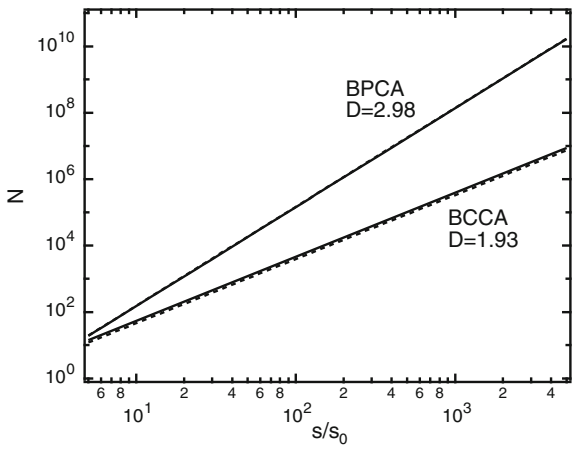


Fig. 3 The number N of constituent grains in fractal agglomerates formed under coagulation growth of the BPCA and the BCCA processes as a function of its characteristic radius normalized to the radius of constituent grains. *Solid line* $\log k_0 = -0.576D + 0.915$ (Kimura et al. 1997); *dashed line* $\log k_0 = -0.5D + 0.7$ (Mukai et al. 1992). From Kimura et al. (1997)

$$a_c = \sqrt{\frac{5}{3}} a_g, \tag{4}$$

which is known to well describe an apparent radius of agglomerates. In reality, constituent grains are not identical, but a nucleation theory of pristine condensates in a protoplanetary disk predicts a size distribution of pristine grains with a single peak in an extremely narrow size range (Yamamoto and Hasegawa 1977). In addition, the major constituents of CP IDPs called GEMS (glass with embedded metal and sulfides) grains are typically limited to $a_0 = 0.05-0.25\,\mu m$ (Keller and Messenger 2011). Therefore, the radius of constituent grains in an agglomerate does not seem to vary considerably as far as primitive dust particles in planetary systems are concerned. Consequently, it is popular practice to use agglomerates of identical constituent grains in an approximation of primitive dust particles in planetary systems. One should, however, keep in mind that the size distribution of constituent grains influences light-scattering and thermal-emission properties of agglomerates to a certain extent.

In popular open-source light-scattering codes available to date, the size parameter of an agglomerate is defined as $x_v = 2\pi a_v/\lambda$ where a_v is the radius of volume equivalent sphere and λ is the wavelength of interest. Note that both a_v and x_v are quantities that are independent of the coagulation process, because $a_v = N^{1/3} a_0$. As a result, an apparent radius of agglomerates cannot be described by a_v at all, but a_v is a useful quantity if the mass or volume of the agglomerates is of interest.

Figure 4 illustrates an example of fractal agglomerates consisting of 2^{13} spherical grains formed under hit-and-stick coagulation processes along ballistic

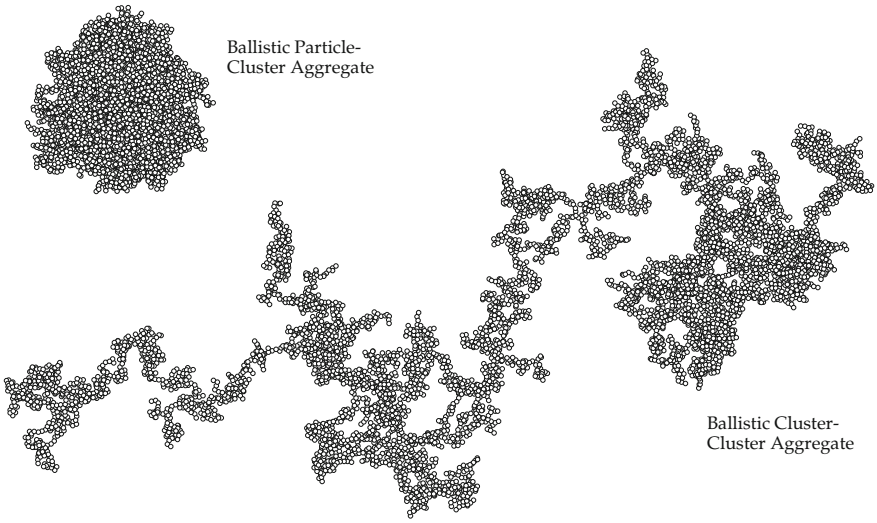


Fig. 4 Two realizations of dust agglomerates that formed by the BPCA and the BCCA processes. The coagulation growth was conducted to hit and stick on their respective contacts, while 2^{13} constituent grains were assumed to be identical spheres. From Kimura (2001)

trajectories. The agglomerates were grown by either the ballistic cluster–cluster agglomeration (BCCA) or the ballistic particle–cluster agglomeration (BPCA) of constituent grains. The fractal dimensions of BCCA and BPCA particles have been determined numerically to be approximately $D \approx 2$ and 3, respectively (Meakin 1984). Agglomerates grown by the diffusion limited agglomeration (DLA) process are also fractals with $D \approx 2.5$ if their constituent grains are assumed to hit and stick on contact (Meakin 1984). Although the DLA process is not relevant to coagulation growth in a protoplanetary disk, high-velocity collisions of BCCA particles result in a fractal dimension $D \approx 2.5$ at their maximum compression as shown in Fig. 5 (Wada et al. 2008). DLA particles could, therefore, be used to represent highly compressed BCCA particles if the fractal dimension plays a vital role in the determination of their light-scattering and thermal-emission properties.

It is worth noting that the configuration of constituent grains in an agglomerate is far from well known for primitive dust particles in planetary systems, although the morphology of CP IDPs is sometimes considered typical of the particles. We argue that CP IDPs represent the most compact structure of primitive dust particles in planetary systems, since there is natural selection with respect to the morphology. If CP IDPs are primitive dust particles released from comets or asteroids, then they must have been injected into a bound orbit around the Sun. It is well known that dust particles in orbit around a star could spiral to the central star by a relativistic effect, called the Poynting–

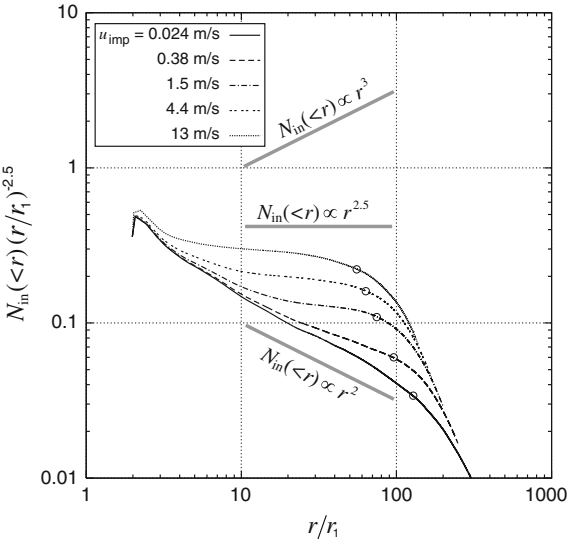


Fig. 5 The number of constituent grains within a distance r from the center of an agglomerate as a function of the distance. Here the number is normalized by the distance to the power of 2.5 and the distance is normalized by the radius of constituent grains so that the fractal dimension $D = 2.0$, 2.5, and 3.0 gives a slope of −0.5, 0, and 0.5, respectively. The agglomerates were produced by numerical simulations on mutual collisions of identical BCCA particles consisting of 2^{13} grains at various collision velocities based on contact mechanics. From Wada et al. (2008)

Robertson effect (Robertson 1937; Wyatt and Whipple 1950; Kimura et al. 2002a). However, the fluffier the morphology of dust agglomerates is, the more efficiently the stellar radiation pressure blows out the agglomerates into the interstellar medium. As a result, it is dynamically infeasible that dust agglomerates with a more fluffy structure stay in a bound orbit around the Sun. Therefore, CP IDPs must be so compact that they could stay in a bound orbit after their release from their parent bodies and reached the Earth by the Poynting–Robertson effect.

In 1990s, the radius of constituent grains in primitive dust agglomerates was often assumed to be $a_0 \sim 0.01\,\mu\mathrm{m}$, simply to alleviate computational limitations (e.g., Mukai et al. 1992; Kozasa et al. 1993; Kimura et al. 1997). However, laboratory studies on the morphology of CP IDPs as well as theoretical studies on the optical properties of cometary dust and dust in debris disks indicate that the radius of constituent grains is on the order of submicrometers (Brownlee 1985; Kimura et al. 2003c; Graham et al. 2007). In addition, ameboid olivine aggregates in the Allende meteorite and aggregate-type particles in the samples of asteroid 25143 Itokawa are characterized by constituent grains whose radii are on the order of micrometers (Grossman and Steele 1976; Yada et al. 2014). Therefore, we regard $a_0 \sim 0.1-1\,\mu\mathrm{m}$ as a stringent constraint on a plausible model of dust agglomerates in planetary systems.

3 Light-Scattering Techniques for Dust Agglomerates

3.1 T-Matrix Method and Generalized Multiparticle Mie Solution

The superposition T-matrix method (TMM) is robust when numerical simulation of light scattering is performed for a cluster of spheres particularly in random orientations (Mackowski and Mishchenko 1996). Although the advantage of the TMM is its efficiency in analytical averaging of scattering characteristic over particle orientations, huge random-access memory (RAM) is a requisite for the publicly available code `scsmtm1` developed by Mackowski and Mishchenko (1996) when used for large agglomerates. Therefore, Okada (2008) proposed to use another code `scsmfo1b` developed by Mackowski and Mishchenko (1996), which gives a solution in a fixed orientation, and to numerically perform orientation averaging with a quasi-Monte Carlo (QMC) method. Because the requisite RAM is smaller in `scsmfo1b` than `scsmtm1`, the QMC method provides an opportunity of dealing with large agglomerates of submicron constituent grains. Penttilä and Lumme (2011) recommend the optimal cubature on the sphere (OCoS) method to perform fast numerical orientation averaging along with the `scsmfo1b` code. The new publicly available `MSTM` (Multi Sphere T-Matrix) code in Fortran-90 is intended to replace the `scsmtm1` and `scsmfo1b` codes in Fortran-77, so that the `MSTM` code optimally uses the memory and processor resources (Mackowski and Mishchenko 2011).

Similar to the superposition TMM, the GMM is an exact method for computing light-scattering properties of dust agglomerates consisting of spherical grains (Xu 1995; Xu and Gustafson 2001). A T-matrix formulation of the GMM provides fast computing of light scattering by an agglomerate of spherical constituent grains in random orientation (Xu 2003). In 2000s, a prominent feature of GMM was the ability of providing a solution to light-scattering problems for an agglomerate consisting of spherical core–mantle grains with the use of the publicly available code `gmm02TrA` (Xu and Khlebtsov 2003). Another code `gmm02f` for a fixed orientation is expected to deal with larger agglomerates of core–mantle spheres than `gmm02TrA` for random orientation, but we are aware that the `gmm02f` code contains a bug, which has not been resolved to date. In 2010s, all the features of the `gmm02TrA` code can be provided by the `MSTM` code of the TMM along with the ability of parallel computing.

3.2 Discrete Dipole Approximation

The discrete dipole approximation (DDA) is a flexible and powerful method for computing light-scattering properties of arbitrary shaped dust particles (Purcell and Pennypacker 1973; Draine and Flatau 1994). In the DDA, dust particles are represented by an array of point electric dipoles and thus the determination of dipole polarizability plays a key role in determining the responses of the dipoles to electromagnetic waves (Draine and Goodman 1993; Collinge and Draine 2004). Hage and Greenberg (1990) demonstrated that the dipole polarizability determined by the digitized Green function/volume integral equation formulation (DGF/VIEF) gives more accurate results than the Clausius–Mossotti relation with the radiative reaction correction employed by Draine (1988). The lattice dispersion relation (LDR) has been implemented as a default prescription for the dipole polarizabilities in the publicly available code `DDSCAT` developed by Draine and Flatau (1994). Okamoto (1995) demonstrated that the a_1-term method is superior to the LDR method for agglomerates of spherical constituent grains, since the former takes into account the boundary condition of spheres (see also Okamoto and Xu 1998). Therefore, one should keep in mind that the choice of polarizability prescription is not a minor issue for the accuracy of numerical results computed by the DDA. Another open-source code for the DDA is `ADDA` that has been intended to run on a multiprocessor memory-distributed system (Yurkin and Hoekstra 2011). The advantage of the DDA over the TMM is the ability to deal with not only arbitrary shapes of constituent grains in an agglomerate but also a contact spot formed by adhesion between elastic constituent grains.

The DDA has been formulated in a way that it provides a solution to a fixed orientation of a particle with respect to the direction of incident electromagnetic wave. As a result, it is common practice to numerically perform orientation averaging of scattering properties by solving the equations for each orientation sequentially. However, for highly fluffy agglomerates, the solution is very sensitive

to their orientations and thus orientation averaging in the DDA is time consuming. Singham et al. (1986) provide an exact analytic expression for the orientational average of scattering matrix elements in the DDA, but the implementation of the analytic averaging in modern DDA formulations is not straightforward (Yurkin and Hoekstra 2007). Mackowski (2002) proposed to use the DDA for computing the T-matrix of arbitrary shaped particles and then apply analytic formulas to the T-matrix for orientation averaging. This is most likely the best method for orientation averaging of scattering properties in the framework of the DDA, but this method has not been applied to dust agglomerates in planetary systems.

3.3 *Effective Medium Approximations*

An effective medium approximation (EMA) is computer friendly and thus has been a popular method to estimate light-scattering and thermal-emission properties of dust agglomerates in combination with Mie theory. Among EMAs, the Maxwell Garnett mixing rule and the Bruggeman mixing rule are the two most popular methods in astronomy. In the Maxwell Garnett mixing rule, dust agglomerates of small constituent grains are represented by a hypothetical inhomogeneous sphere with inclusions of constituent grains in the matrix of vacuum. In the Bruggeman mixing rule, agglomerates of small constituent grains are assumed to be a hypothetical inhomogeneous sphere consisting of constituent grains and "vacuum" grains.

It is important for EMAs to properly define the radius of a hypothetical porous sphere and its porosity, since the definitions of these quantities are not unique. Mukai et al. (1992) proposed to use the characteristic radius a_c of an agglomerate and the porosity defined by

$$p = 1 - \frac{V}{V_\mathrm{c}}, \tag{5}$$

where V and V_c are the total volume of constituent grains and the volume of a hypothetical sphere with a radius a_c. It turned out that these definitions enable EMAs to well reproduce the radiation pressure on agglomerates computed by the DDA, as far as agglomerates with $a_0 = 0.01$ μm are concerned (Mukai et al. 1992; Kimura et al. 2002a). Note that Eq. (5) can be written as

$$p = 1 - k_0 \left(\frac{a_\mathrm{c}}{a_0} \right)^{D-3}, \tag{6}$$

in case of fractal agglomerates. This clarifies that the porosity for agglomerates with $D = 3$ is independent of a_c and the porosity for agglomerates with $D < 3$ increases with a_c (see Fig. 6). Therefore, as illustrated in Fig. 4, fractal agglomerates with $D \approx 3$ are relatively compact compared to those with $D \approx 2$.

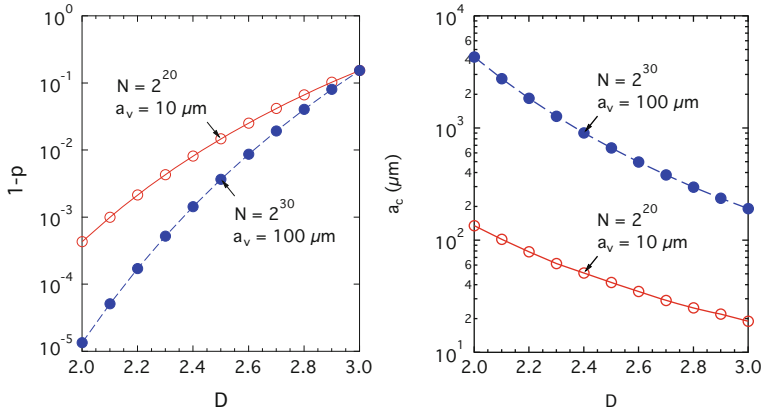


Fig. 6 The volume fraction $1-p$ and the characteristic radius a_c of fractal agglomerates consisting of 0.1 μm-radius spherical grains as a function of fractal dimension D. *Open circles with a solid line* and *filled circles with a dashed line* are the corresponding values at $N = 2^{20}$ ($a_v = 10\,\mu m$) and $N = 2^{30}$ ($a_v = 100\,\mu m$), respectively. Reproduced from Table 1 of Kimura et al. (2009)

4 Light Scattering by Dust Agglomerates

4.1 Single Scattering

Most of light-scattering phenomena associated with primitive dust particles in planetary systems are well described by scattering of stellar radiation on single particles. The geometric albedo of dust particles is one of the quantities that can be derived from both numerical simulations and astronomical observations of light scattering by the particles. It is defined by the ratio of stellar radiation scattered by the particles at zero phase angle (exact backscattering) to that scattered by a white Lambertian flat disk of the same geometric cross section. The geometric albedos of micrometeoroids, dust particles in the zodiacal cloud, and the coma of comet 1P/Halley have been determined to be <0.1, 0.06, 0.04, respectively (Hanner 1980; Ishiguro et al. 2013; Lamy et al. 1989). These values reveal visibly very dark appearances of primitive dust particles at least in the Solar System, implying that the particles are dominated by carbonaceous materials. Hanner et al. (1981) extended the definition of geometric albedo so that the brightness of fluffy agglomerates at any scattering angle is normalized to the geometric albedo at backscattering. In this way, the geometric albedo A_p of a particle at a wavelength of λ is proportional to the $(1,1)$ element of the Mueller matrix, S_{11}, as $A_p = S_{11}\lambda^2/(4\pi G)$ where G is the geometric cross section of the particle. The degree of linear polarization is another quantity that gives a direct comparison between numerical simulations and observations.

With the help of the DDA, West (1991) has pioneered numerical simulation of light scattering by DLA ($D \approx 2.5$) particles consisting of submicrometer-sized grains. It is most likely that he employed the Clausius–Mossotti relation with the radiative reaction correction given by Draine (1988) to determine the polarizability of each dipole, since he used the early version of the DDSCAT code. The number of the constituent grains was set to either $N = 170$ or 8 and the size parameter x_0 of the grains, which is defined as $x_0 = 2\pi a_0/\lambda$, ranged from $x_0 = 0.19$ to 0.57 for $N = 170$ and $x_0 = 0.6-1.8$ for $N = 8$. The intensity of light scattered by agglomerates shows an enhancement toward forward scattering that is characteristic of large particles, but agglomerates do not display undulations in its angular dependence as opposed to spheres of similar geometrical cross sections. Interestingly, the angular dependence of linear polarization is well characterized by the size of constituent grains in agglomerates, rather than the overall size of the agglomerates.

Kozasa et al. (1992), (1993) intensively studied light-scattering properties of BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles consisting of tiny constituent grains with $a_0 = 0.01$ μm. They implemented DDA computations using the DDSCAT code (ver. 4a) along with the DGF/VIEF method for the determination of dipole polarizability. Kozasa et al. (1992) presented the dependences of absorption and scattering cross sections and asymmetry parameter on wavelength, while Kozasa et al. (1993) considered the dependences of intensity and linear polarization on scattering angle for silicate and magnetite agglomerates at various numbers of the constituent grains up to $N = 4096$. They also investigated the validities of empirical formulas as well as the Maxwell Garnett mixing rule for the light-scattering properties of the agglomerates. Although the size of their constituent grains at a wavelength $\lambda = 0.6\,\mu$m lies in the Rayleigh scattering regime ($x_0 = 0.1$), their results for BPCA ($D \approx 3$) particles revealed that the maximum degree of linear polarization decreases with the number of constituent grains. Consequently, we may expect that the degree of linear polarization strongly depends on the size of constituent grains, but to a lesser degree the apparent size of agglomerates.

Lumme et al. (1997) provided a survey of light-scattering properties with agglomerates of submicron spherical constituent grains relevant to primitive dust particles in planetary systems. They accomplished the survey using their own DDA code with icy and silicate agglomerates consisting of $N = 200$ spherical grains. The size parameter x_0 of constituent grains was assumed to be $x_0 = 1.2$ or 1.9, which corresponds to submicron constituent grains in the visible wavelength range. The arrangement of the spheres in an agglomerate follows either the DLA process or a removal of spheres from a maximum packed cluster of spheres. The a_1-term method was applied to determine the dipole polarizability, although the polarizability was truncated to the first three terms in the expansion of the a_1-term with respect to the size parameter of the dipoles. Their results show no clear difference in the intensity between ice and silicate agglomerates, but otherwise they provide a supplement to the results by West (1991). They suggested that the DDA and the TMM should be combined with ray-tracing techniques if one considers a broad range of agglomerate sizes.

Xing and Hanner (1997) considered agglomerates of tetrahedral constituent grains as well as those of spherical constituent grains in order to study the effect of grain's shapes on the angular dependences of intensity and polarization at $\lambda = 0.6\,\mu\text{m}$. Their numerical simulations were performed using the DDSCAT code (ver. 4b) with the LDR method for the determination of dipole polarizabilities. Agglomerates of glassy carbon were assumed to consist of N = 4 or 10 grains with a radius of $a_0 = 0.25$ or $0.5\,\mu\text{m}$, which are either touching, overlapping, or separated, while silicate agglomerates were assumed to consist of 10 touching tetrahedral grains with $a_0 = 0.25\,\mu\text{m}$. In comparison to the degree of linear polarization, the angular dependence of intensity showed less variations with the size and shape of the agglomerates and the shape of the grains. Although they concluded that a mixture of carbon and silicate agglomerates represents cometary dust, the presence of ripples in the angular dependences of intensity and polarization has not been detected by observations to date.

Levasseur-Regourd et al. (1997) computed the angular dependences of intensity and linear polarization for BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles of spherical constituent grains at $\lambda = 0.62\,\mu\text{m}$. They employed the DDSCAT code (ver. 4b) with presumably the LDR method to compute the light-scattering properties of the agglomerates. The number of constituent grains was fixed to $N = 512$, but the size of the agglomerates lay in the range from $x_\text{v} = 0.90$ to 6.02, in other words, $x_0 = 0.11{-}0.75$. Because the angular dependences of intensity and polarization for BPCA particles exhibit ripples at $x_\text{v} > 2$ regardless of the refractive index, they concluded that primitive dust particles in the Solar System are agglomerates with $D \sim 2$. On the one hand, we are unable to agree with their conclusion, since their results have not been reproduced by any of the later studies with BPCA particles at $x_\text{v} > 2$ (cf. Kimura 2001; Kimura et al. 2003c, 2006; Bertini et al. 2007; Kolokolova and Mackowski 2012). On the other hand, we notice that compact agglomerates at $x_\text{v} > 2$ may exhibit ripples in their angular dependences of intensity and polarization if their constituent grains are located on a periodic lattice or distributed in a more symmetric manner (cf. Petrova et al. 2000, 2001a; Kolokolova and Mackowski 2012). It is most likely that their computations were performed with fast Fourier transform methods that require dipoles to be located on a periodic lattice (see Draine and Flatau 1994). Because constituent grains were replaced by single dipoles in their computations, we cannot help wondering whether the ripples resulted from the displacement of constituent grains on a periodic lattice.

To better understand optical observations of zodiacal light at small phase angles, Nakamura and Okamoto (1999) studied the angular dependences of intensity and polarization for BPCA ($D \approx 3$) particles consisting of spherical silicate grains. They used the DDSCAT code (ver. 4a) with the a_1-term method by assuming $a_0 = 0.03\,\mu\text{m}$ at $\lambda = 0.5\,\mu\text{m}$ so that $x_0 < 1$. Their results revealed that the gegenschein and the negative branch of linear polarization are insensitive to the number of constituent grains, which lay in the range from $N = 8000$ ($a_\text{v} = 0.6\,\mu\text{m}$) to $N = 27,000$ ($a_\text{v} = 0.9\,\mu\text{m}$). They attributed quantitative disagreements between their numerical results and zodiacal light observations to the size and shape of the constituent grains. However, for a quantitative discussion on a model of zodiacal light, one should integrate light

scattered by agglomerates along the line of sight. Furthermore, it is clear that large agglomerates ($a_v \gg 9\,\mu m$) are required to model the strong forward scattering of zodiacal light.

Kimura (2001) investigated intensively how the angular dependences of intensity and linear polarization for fractal agglomerates of spherical constituent grains depend on the grain size at a wavelength $\lambda = 0.6\,\mu m$. The radius of constituent grains lay in the range of $a_0 = 0.01-0.15\,\mu m$ ($x_0 = 0.10-1.57$) and the configuration of the grains was determined by either the BPCA ($D \approx 3$) or the BCCA ($D \approx 2$) process. The `DDSCAT` code (ver. 4a) with the a_1-term method and the `scsmtm1` code were applied to agglomerates of up to 8192 grains with $a_0 = 0.01\,\mu m$, 512, grains with $a_0 = 0.07\,\mu m$, and 256 grains with $a_0 = 0.15\,\mu m$. It turned out that the morphology of agglomerates is of importance for the smallest grains with $a_0 = 0.01\,\mu m$, but otherwise the angular dependences of intensity and linear polarization are independent of coagulation process. A comparison of the results between silicate and carbon agglomerates indicates that the composition of the constituent grains strongly affects the optical properties of agglomerates with $a_0 = 0.15\,\mu m$ ($x_0 = 1.57$) as opposed to the comparison between ice and silicate agglomerates studied by Lumme et al. (1997). Interestingly, the degree of linear polarization becomes negative at backscattering not only for silicate agglomerates but also for carbon agglomerates if $x_0 \sim 1$ and $x_v \gg 1$ (see Fig. 7). This result shed new light on the presence of the negative polarization branch at small phase angles for primitive dust particles whose geometric albedo is very small.

Petrova et al. (2000), (2001a) used the `scsmtm1` code to study the optical properties of silicate agglomerates consisting of spherical grains that were arranged in a tetrahedral or cubic lattice. The size parameter of the constituent grains lies in the range of $x_0 = 0.7-2.5$, which corresponds to $a_0 = 0.07-0.25\,\mu m$ at a wavelength of $\lambda = 0.63\,\mu m$ and $a_0 = 0.06-0.21\,\mu m$ at $\lambda = 0.54\,\mu m$. A special feature of their work is that a size distribution of the agglomerates was taken into consideration, although their agglomerates were composed of only 1–43 grains. They claim that the intensity and the polarization of their agglomerates and cometary dust are similar with respect to their angular dependences, if the grain radius lies in the range of $x_0 = 1.3-1.65$. As the authors noticed, however, their results on the intensity and the polarization do not provide conclusive evidence for their wavelength dependences, owing to a lack of large agglomerates in their models.

Petrova et al. (2004) extended their study on angular dependences of intensity and polarization to larger silicate agglomerates of 12–150 constituent grains. They used BPCA and DLA particles consisting of spherical grains with $a_0 \approx 0.1\,\mu m$ and considered a power-law size distribution of the particles with the power of -3. The agglomerates were constructed on the assumption that the constituent grains hit and stick on contact only if the grains have two or more contacts. Although the agglomerates were still small, they were characterized by a fractal dimension $D \approx 2$–3 and a porosity $p \approx 0.95-0.63$. Their results show that the degree of the maximum polarization is higher and the scattering angle of the maximum polarization is larger than observed for comets. They attributed the discrepancies between their model results and observations to insufficiency of the number of

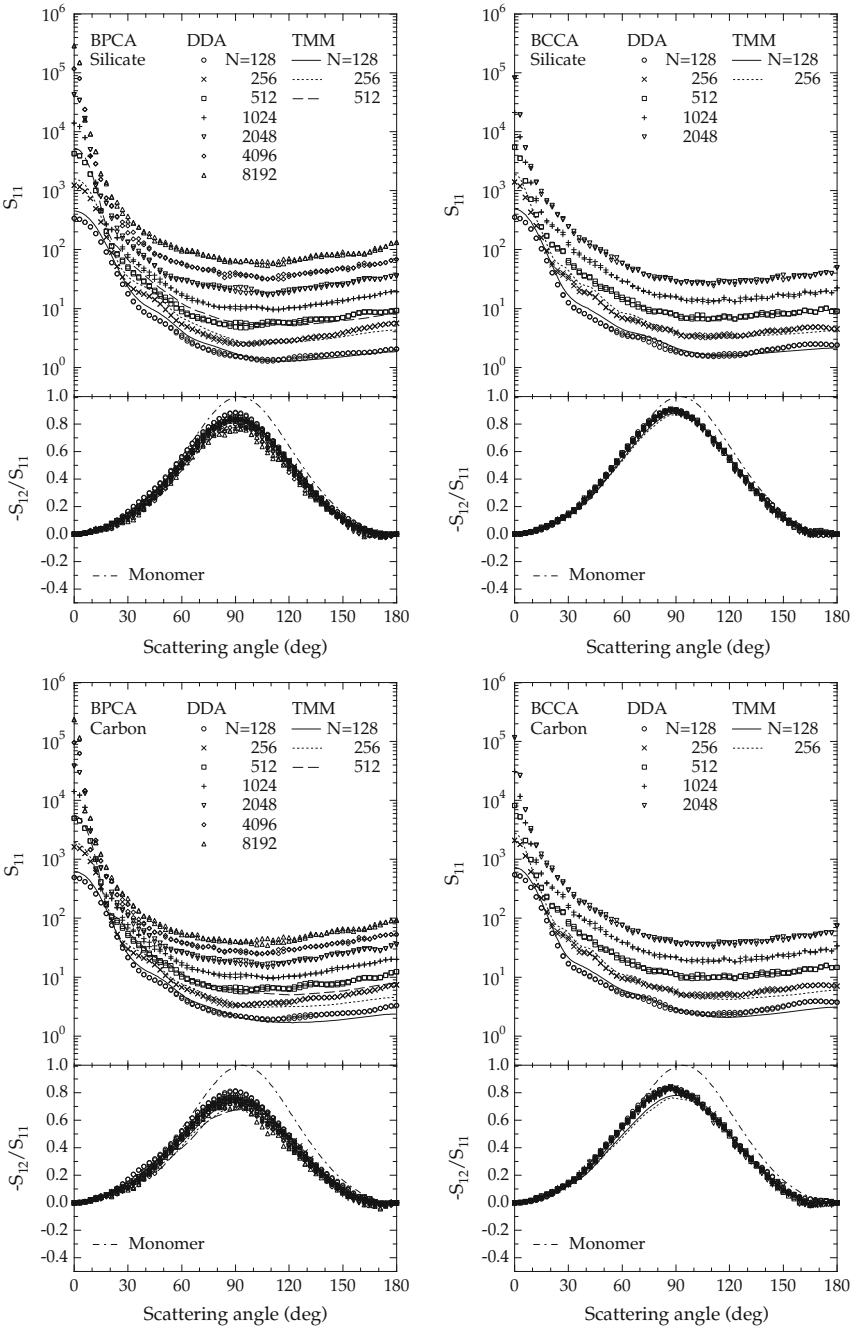


Fig. 7 The dependences of intensity S_{11} and linear polarization $-S_{12}/S_{11}$ on scattering angle at a wavelength $\lambda = 0.6\,\mu\text{m}$ for BPCA particles (*left*) and BCCA particles (*right*) consisting of silicate grains (*upper panels*) and carbon grains (*lower panels*) with $a_0 = 0.07\,\mu\text{m}$. Also plotted as *dash-dotted lines* are the $-S_{12}/S_{11}$ values for the constituent grains. From Kimura (2001)

constituent grains, but part of the discrepancies most likely resulted from low refractive indices assumed in their paper.

It is worth noting that the dependences of geometric albedo and linear polarization on scattering angle and wavelength for cometary dust show common characteristics among a variety of comets (Gustafson and Kolokolova 1999). Kimura et al. (2003c) succeeded for the first time in qualitatively reproducing all the common characteristics of cometary dust simultaneously using BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles of optically dark submicron constituent grains (see Fig. 8). The authors first determined the average refractive indices of the constituent grains by applying the Maxwell Garnett mixing rule to a mixture of silicate, iron, organic refractory, and amorphous carbon. The volume fractions of these substances were derived from the elemental abundances of dust in comet 1P/Halley that were measured in situ by PUMA-1 onboard VeGa 1 (see Jessberger et al. 1988, for the elemental abundances). The average refractive indices were then used for their numerical computations with the `scsmtm1` code for agglomerates of identical spherical constituent grains with $a_0 = 0.1\,\mu\text{m}$. Their results did not show a clear dependence on the structure of agglomerates between the BPCA and BCCA processes nor the number of constituent grains. Therefore, their success in reproducing the common characteristics can be attributed to the use of optically dark material, which is expected from the composition of cometary dust (see also Mann et al. 2004; Kolokolova et al. 2005). In conclusion, the presence of common

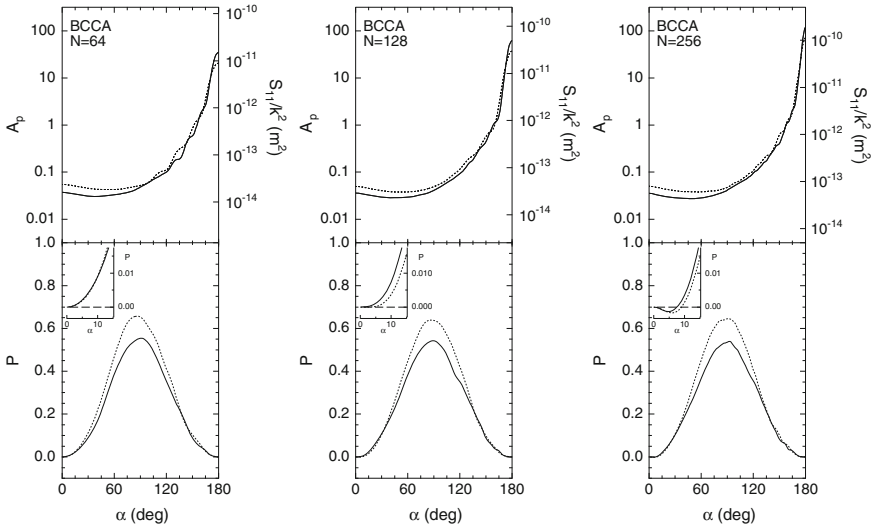


Fig. 8 The geometric albedo A_p and linear polarization P of BCCA ($D \approx 2$) particles consisting of monodisperse spherical grains as a function of phase angle α. The constituent grains have a radius $a_0 = 0.1\,\mu\text{m}$ and a composition inferred from the elemental abundances measured for comet 1P/Halley. *Solid lines* and *dotted ones* are the corresponding values at a wavelength $\lambda = 0.45\,\mu\text{m}$ and $\lambda = 0.60\,\mu\text{m}$, respectively. From Kimura et al. (2003c)

light-scattering characteristics in cometary dust reflects the fact that comets formed out of the same protoplanetary disk materials with the solar composition.

Kimura et al. (2006) have thoroughly explored which size, number, composition, and configuration of constituent grains reproduce the common light-scattering characteristics of cometary dust. The authors applied the `scsmtm1` code to numerical simulation of light scattering by BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles of up to 256 spherical constituent grains. Their numerical results were used to place constraints on the size and composition of constituent grains, confirming the successful model of Kimura et al. (2003c) based on the properties of CP IDPs and dust in comet 1P/Halley. The contribution of amorphous carbon to the results is so significant that the common light-scattering characteristics of cometary dust could be reproduced even by agglomerates without organic materials, in particular, agglomerates of amorphous carbon. It turned out that the size and composition of constituent grains play a crucial role in the determination of optical properties. Although the size and composition of constituent grains are well constrained in the framework of the model, quantitative fits to the observed angular dependence of polarization seem to require a larger number and non-spherical shape of constituent grains (cf. Kimura and Mann 2004).

A similar numerical approach to constraining the morphology and composition of cometary dust was taken by Bertini et al. (2007). They used the `DDSCAT` code (ver. 5a10) presumably with the LDR method to study light scattering by BPCA particles of up to 100 spherical particles with $a_0 = 0.13-0.16\,\mu\text{m}$ at $\lambda = 0.535$, 0.6274, and 1.5 μm. They assumed the agglomerates to be composed of silicates, organic materials, or a mixture of silicates and organic materials with equal mass. Unfortunately, they failed to find a solution for the morphology and composition of agglomerates that simultaneously reproduces all the common characteristics of cometary dust. Nevertheless, they suggest that two distinct types of linear polarization among different comets suggested by Levasseur-Regourd et al. (1996) arise from the difference in the radius of constituent grains. However, as discussed in Kolokolova et al. (2007), the presence of two polarimetric classes might be artifact due to a contribution of gas emission to the continuum measured with broadband filters.

Using the `scsmtm1` code, Kolokolova and Kimura (2010) modeled cometary dust with two types of BCCA particles consisting of 256 identical spherical grains with $a_0 = 0.1\,\mu\text{m}$, along with a multi-shaped, polydisperse mixture of spheroids. The authors considered not only the elemental abundances of dust in comet 1P/Halley, but also the mineralogical classification of the dust (see Fomenkova 1999, for the classification). In the model, the agglomerates are composed of either organic refractory material alone or a mixture of silicate, metal, organic refractory, and amorphous carbon, and the spheroids are composed of silicate alone. The three types of particles were intended to model organic-rich, silicate-poor particles, organic-poor, silicate-rich particles, and compositionally intermediate particles. The results have shown that the model is successful in reproducing all the common characteristics of cometary dust not only qualitatively, but also quantitatively (see Fig. 9). This might indicate that a mixture of compositionally different particles is a

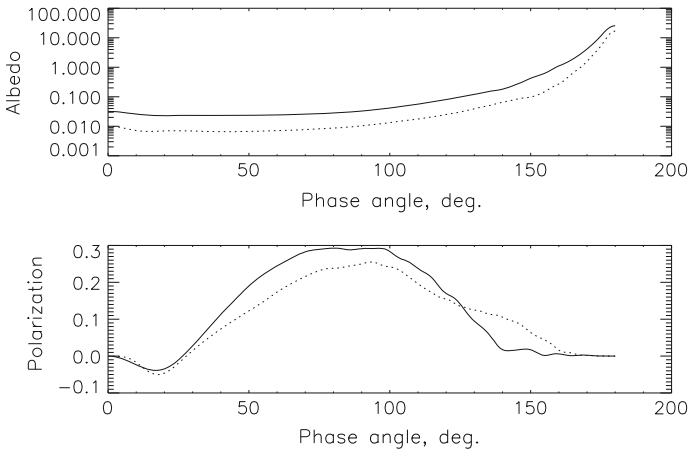


Fig. 9 The dependences of geometric albedo A_p and linear polarization P on phase angle α for a mixture of BCCA ($D \approx 2$) particles consisting of $N = 256$ spherical grains with $a_0 = 0.1\,\mu m$ and compact spheroidal particles with a power-law size distribution. *Dotted lines* and *solid ones* are the corresponding values at wavelengths $\lambda = 0.45\,\mu m$ and $\lambda = 0.60\,\mu m$, respectively. The agglomerates are either an admixture of silicates, metals, and carbonaceous materials, or pure organic refractory material, while the spheroids are pure silicates. From Kolokolova and Kimura (2010)

key to the common characteristics of cometary dust, instead of large agglomerates suggested by Kimura et al. (2006).

Moreno et al. (2007) applied the DDSCAT code (ver. 6.0) with the LDR method to simulate light scattering by cometary dust with DLA particles ($D \approx 2.5$) of up to 256 constituent grains. They assumed that either constituent grains are cubes with the side length of 0.15 μm or they are spheres with $a_0 = 0.075\,\mu m$. Using the same refractive indices as Kimura et al. (2003c), their results show that the maximum degree of linear polarization for agglomerates of cubes is lower than that for agglomerates of spheres, while the maximum takes place at a smaller scattering angle for cubic constituent grains than spherical ones. However, it should be noted that the diameter of spherical constituent grains equals the side length of cubic constituent grains, implying that the volume of the spheres is smaller than that of the cubes. Therefore, it is not clear whether the difference in the results between cubic constituent grains and spherical ones arises from the shape or the size of the grains.

Lumme and Penttilä (2011) employed not only agglomerates of spherical constituent grains but also those of Gaussian random constituent grains with the same volume as the spherical grains. They used the scsmfo1b code with the OCoS method to the agglomerates of spherical grains and the ADDA code presumably with the LDR method to the agglomerates of Gaussian random grains. The morphology of agglomerates was determined either by the BPCA process with controls on the ballistic trajectories or by the BCCA process with an assemblage of 16 BPCA particles at a single point. The radius, number, and composition of constituent grains

lay in the range of $0.05 \leq a_0 \leq 0.25$ μm ($0.5 \leq x_0 \leq 2.5$), $128 \leq N \leq 512$, and ice to silicate, respectively. Both the constituent grains and the agglomerates were assumed to either be monodisperse or have a size distribution. A comparison of the results between agglomerates of spherical constituent grains and those of Gaussian random constituent grains shows similar angular dependences of intensity and polarization, although there are noticeable differences in their magnitudes. They claim the presence of a strong correlation between the real part and the imaginary part of refractive index that results in almost the same angular dependences of intensity and polarization. If this correlation is universal, then any attempts to fit observational data without the geometric albedo will not provide a unique solution.

Zubko et al. (2011), (2012), (2013), (2014) consider cometary dust as agglomerates of irregularly shaped polydisperse constituent grains, along with a power-law size distribution of the agglomerates (see also Zubko 2012, 2013). They apply the DDA with the LDR method to solve light-scattering problems using their own code ZDD, instead of the publicly available DDSCAT and ADDA codes (Penttilä et al. 2007; Zubko et al. 2010). The main problem of their approach is how the size distribution of the agglomerates was achieved, since they used the same agglomerates to scale larger and smaller sizes of agglomerates. Namely, the larger is the agglomerate, the larger are its "constituent grains," but in all other respects all the agglomerates are identical to one another. In reality, larger agglomerates have beyond a shadow of doubt a larger number of constituent grains, not the larger size of the grains.

Lasue and Levasseur-Regourd (2006) presented their numerical simulations on light scattering by BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles of up to 128 spheres or prolate spheroids with the axis ratio of two. The constituent grains have a radius of $a_0 = 0.1$ μm and consist of pure silicate, pure organics, or a mixture of silicate in the core and organics in the mantle. Their computational results with the DDSCAT code (ver. 5a10) presumably with the LDR method showed a great similarity in the optical properties of agglomerates between the organic-coated silicate constituent grains and the pure organic ones. The paper also provides a model of fractal agglomerates with a size distribution that fits the observational data on the angular dependence of linear polarization for comet C/1995 O1 (Hale–Bopp). Despite the best fit to the observed polarization in the range of phase angles from 0° to 50°, the geometric albedo of the agglomerates seems to be higher than observed.

Lasue et al. (2009) extended their light-scattering modeling of cometary dust using agglomerates of up to $N = 256$ constituent grains of $a_0 = 0.1$ μm and fitting to the polarimetric data for both comets C/1995 O1 (Hale–Bopp) and 1P/Halley. They used the DDSCAT code (ver. 6.1) presumably with the LDR method for their numerical simulations with agglomerates of spherical or spheroidal constituent grains. Although they obtained a good fit to the angular dependence of linear polarization for comet Hale–Bopp using a mixture of agglomerates and compact spheroids, they failed to reach the observed values of negative polarization for comet 1P/Halley. It is unfortunate that this study provides no clue about the geometric albedo nor the color of their agglomerates and that the size of the constituent grains is scaled to represent the size distribution of the agglomerates.

Golimowski et al. (2006) presented multiband coronagraphic images of the debris disk around the A-type star β Pictoris. They modeled the observed red colors of the disk using light-scattering properties of porous agglomerates computed by Wolff et al. (1998) and Voshchinnikov et al. (2005). On the one hand, Wolff et al. (1998) used the DDSCAT code (ver. 4c) with the LDR method for agglomerates consisting of pure silicate grains encased in a spherical volume with 60 % of vacuum. On the other hand, Voshchinnikov et al. (2005) used the DDSCAT code (ver. 6.0) with the LDR method for agglomerates consisting of silicate grains and graphite grains half and half encased in a spherical volume with 33 or 90 % of vacuum. Golimowski et al. (2006) reconciled model results with the observed color of the disk except for highly porous agglomerates consisting of silicate grains and graphite grains. It should be, however, noted that the observed red color does not refute a predominance of highly porous agglomerates in the disk, because fractal agglomerates of submicron constituent grains with a rocky core and a carbonaceous mantle exhibit red colors (cf. Kimura et al. 2003c, 2006).

Graham et al. (2007) derived the phase function and polarization of dust particles in the debris disk of the young (12 Myr) nearby M-type star AU Microscopii as a function of scattering angle from optical measurements by the Hubble Space Telescope. The scattering-angle dependences of phase function and polarization were compared to light-scattering properties of porous spheres and those of agglomerates computed by Petrova et al. (2000) and Kimura et al. (2006). It turned out that the phase function and the polarization are better reproduced by the optical properties of BPCA particles consisting of 128 silicate grains with $x_0 = 0.9$. The color of silicate BPCA particles is blue in the visible wavelength range, which agrees with the observed color of the AU Mic debris disk (cf. Krist et al. 2005). These results are entirely consistent with a picture that dust particles in the AU Mic disk originate from planetesimals, similar to asteroids in the Solar System.

The debris disk of the nearby young (8 Myr) A-type star HR 4796A appears to be a circumstellar dust ring around 70 au from the central star (Schneider et al. 1999). The ring with a red color in the visible and a gray color in the near-infrared wavelength range led Debes et al. (2008) to argue for the presence of complex organic materials like Titan's tholins. In contrast, Köhler et al. (2008) modeled the visible to near-infrared spectrum of the ring with porous agglomerates of amorphous silicate, amorphous carbon, and water ice. They considered that spherical constituent grains of $a_0 = 0.1\,\mu\text{m}$ coagulated into agglomerates with a porosity $p = 0.73$. They used the so-called Henyey–Greenstein phase function, which is an analytic function specified by the single-scattering albedo and the asymmetry parameter (Henyey and Greenstein 1941). The single-scattering albedo and the asymmetry parameter of the porous agglomerates were computed by the Bruggeman mixing rule. It should be noted that the Bruggeman mixing rule does not properly describe the light-scattering properties of agglomerates, because it ignores the interaction between constituent grains which

plays a vital role in light scattering (cf. Kimura and Mann 2004; Kolokolova and Kimura 2010b).

Shen et al. (2009) studied light-scattering properties of BPCA particles with different degrees of restructuring and applied their results to the AU Mic debris disk and cometary comae. The agglomerates were composed either of silicate alone or a mixture of silicate and graphite in equal volumes, although both the compositions are inconsistent with our current understandings of cometary dust. The number and radius of spherical constituent grains were considered to lie in the range of $N = 32-1024$ and $a_0 = 0.0143-0.16\,\mu\text{m}$. The DDSCAT code (ver. 7.0) was used along with the "modefied" LDR method, in which Gutkowicz-Krusin and Draine (2004) corrected a subtle error in the LDR determined by Draine and Goodman (1993). Contrary to previous studies, the computational results at $\lambda = 0.1-3.981\,\mu\text{m}$ have shown that the size of constituent grains does not play a vital role in the light-scattering properties of agglomerates, if the mass and porosity of agglomerates are fixed. However, the results in Kimura (2001) reveal that light-scattering properties of agglomerates vary with the size of constituent grains, even if the mass and porosity of agglomerates are fixed. Therefore, the agglomerates used in Shen et al. (2009) might be so dense that their light-scattering properties represent those of porous spheres than those of fluffy agglomerates.

Kolokolova and Mackowski (2012) used the MSTM code (ver. 3.0) to intensively study linear polarization of light scattered by agglomerates of $N = 1024$ spherical constituent grains with $a_0 = 0.1\,\mu\text{m}$. They have shown how the porosity and the overall size of agglomerates can be constrained by spectroscopic observations of linear polarization from the visible to the near-infrared wavelength range. Regardless of the wavelength, the refractive index was fixed at the value that was suggested for cometary dust at $\lambda = 0.45\,\mu\text{m}$ by Kimura et al. (2003c). As demonstrated in Fig. 10, misleading numerical results on the degree of linear polarization might arise from an artificial configuration of constituent grains. Consequently, it is essential for a correct understanding of light scattering by primitive dust particles in planetary system to properly model dust agglomerates based on their formation mechanisms.

Videen and Muinonen (2015) applied a radiative transfer technique with coherent backscattering to compute light scattering by sparse agglomerates of identical spherical constituent grains whose positions were chosen randomly and uniformly within a spherical volume. They considered two porosities of 0.94 and 0.97, and assumed the constituent grains to have a size parameter $x_0 = 1.76$ and be nonabsorbing. To our knowledge, this is the first study that was successful in computing light scattering by large agglomerates of $N \sim 3 \times 10^8$ despite the fact that the assumption of nonabsorbing material is inappropriate for primitive dust particles in planetary systems. However, their computations show highly unrealistic phase functions where the forward-scattering peak does not appear for large agglomerates of $N > 10^5$. It is unfortunate that the effects of diffraction, which dominate the forward-scattering region, have not been incorporated in their model.

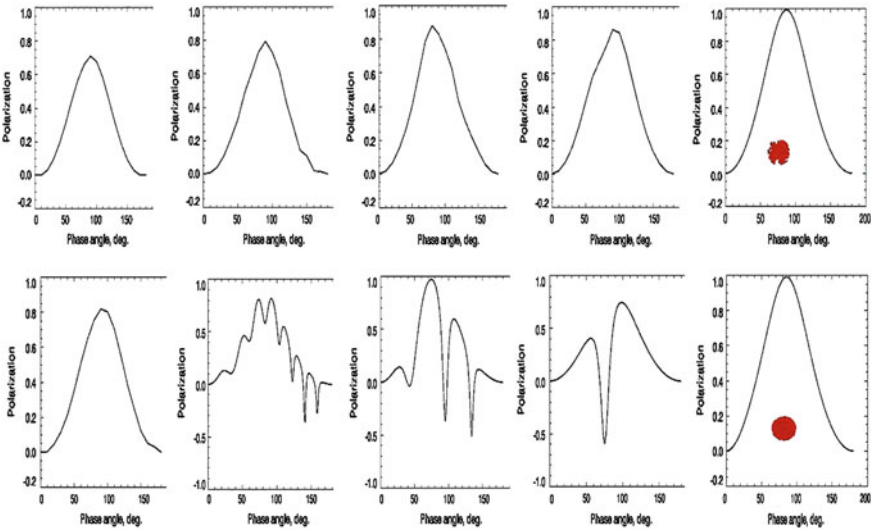


Fig. 10 The degree of linear polarization P as a function of phase angle α for BPCA ($D \approx 3$) particles consisting of $N = 1024$ spherical grains with $a_0 = 0.1\,\mu\text{m}$ (*upper panel*) and equi-dimensional agglomerates consisting of $N = 1024$ spherical grains with $a_0 = 0.1\,\mu\text{m}$ randomly distributed in a spherical volume with equal porosity (*lower panel*) at wavelengths $\lambda = 0.6\,\mu\text{m}$, $1.1\,\mu\text{m}$, $2.2\,\mu\text{m}$, $4.4\,\mu\text{m}$, and $6.0\,\mu\text{m}$ (from *left* to *right*). From Kolokolova and Mackowski (2012)

4.2 Multiple Scattering

Multiple scattering may be of crucial importance in the vicinity of a source region (e.g., a comet nucleus) where dust particles are released from their parent bodies as well as on the surface (e.g., a regolith layer) of their parent bodies. Impacts of micrometeoroids onto the surfaces of asteroids are common phenomena that produce not only regolith particles but also impact craters and ejecta curtains. Dusty ejecta clouds in Saturn's rings produced by impacts of meteoroids onto the rings have been imaged by Cassini's Imaging Science Subsystem (Tiscareno et al. 2013). NASA's Deep Impact mission excavated the surface of comet 9P/Tempel 1 by an artificial impactor to produce an impact crater on the comet as well as an ejecta curtain (A'Hearn et al. 2005). A similar impact experiment and a subsequent imaging observation of an ejecta curtain are planed on asteroid 1999JU3 by the Japanese Hayabusa-2 mission (Arakawa et al. 2013). Recently, radiative transfer computations to model such an ejecta curtain have been performed with assumptions about the phase function, single-scattering albedo, and asymmetry parameter of the ejecta particles (cf. Nagdimunov et al. 2014; Shalima et al. 2015).

Nagdimunov et al. (2014) have succeeded in modeling both an optically thick ejecta plume and its shadow on the surface of comet 9P/Tempel 1 produced and imaged by the Deep Impact mission (see Fig. 11). They considered the ejecta

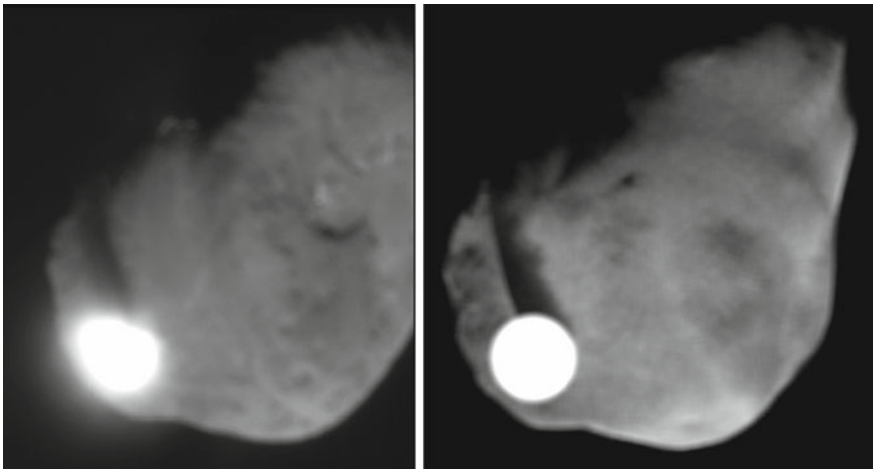


Fig. 11 The real image of the ejecta plume from the surface of comet 9P/Tempel 1 and its shadow on the surface taken by the Deep Impact High-Resolution Instrument (*left*) and the simulated image of the ejecta plume and its shadow computed by radiative transfer computations (*right*). From Nagdimunov et al. (2014)

particles to be porous spheres composed of ice, silicate, amorphous carbon, organic refractory, and "vacuum" materials. The porous spheres were used to simulate light-scattering properties of fluffy agglomerates with a wide size range in the framework of EMAs. They used the Maxwell Garnet mixing rule to compute the single-scattering albedo, asymmetry parameter, and extinction cross section of the agglomerates at a wavelength $\lambda = 0.65\,\mu\text{m}$. However, they adopted the Henyey–Greenstein phase function to determine the scattering angles of photons probabilistically along the path of the photons in their radiative transfer computations. Because the single-scattering albedo and the asymmetry parameter are associated with the phase function of porous agglomerates based on the Maxwell Garnett mixing rule, the use of the Henyey–Greenstein phase function is self-contradictory. It is, therefore, unfortunate that the optical properties of dust agglomerates have not yet been incorporated into available radiative transfer modelings for ejecta curtains in a self-consistent way.

To simulate light scattering by a regolith layer on the surface of asteroids, Petrova et al. (2001b) solved the radiative transfer equation along with the optical properties of agglomerates that were computed by Petrova et al. (2001a). The agglomerates were assumed to have a power-law size distribution and consist of $N = 8{-}43$ spherical grains with $x_0 = 1.50$ or 1.65. The optical depth τ of a plane-parallel regolith layer, in which agglomerates were embedded, lay in the range of $\tau = 0.2{-}50$. Their results have shown that the absolute value of the degree of linear polarization tends to decrease with τ in all scattering angles, although the results at $\tau > 10$ merge into a single curve. It turned out that the negative and positive branches of linear polarization become shallower and higher as the

imaginary part of refractive index decreases. Their results may serve as an explanation for polarimetric observations of asteroids, in which S-type asteroids show a shallower negative branch and a higher positive branch, compared to C-type asteroids.

5 Thermal Emission from Dust Agglomerates

5.1 Spectral Energy Distribution

One of the primary observational data available to constrain the physics of primitive dust particles in planetary systems is its spectral energy distribution (SED) constructed using broadband photometric measurements from mid-infrared to millimeter wavelengths. The SED is dominated by silicate features in the mid-infrared wavelength range from 10 up to 70 μm and provides basic constraints on the dust temperature and optical properties. There is generally no unique solution to SED fitting problems owing in particular to the degeneracies between the optical properties of dust particles and the disk geometry. For example, dust particles in the vicinity of the central star are hotter than the particles of the same size and composition located far from the star, while small opaque particles are hotter than larger ones at the same distance from the star and large transparent particles are hotter than smaller ones at the same distance.

Li and Greenberg (1998) reproduced the 10 μm silicate emission and the SED of the dust disk around β Pic using a model of very porous agglomerates, similar to cometary dust. They applied the Maxwell Garnett mixing rule to the computations of the SED from agglomerates of crystalline silicate constituent grains as well as agglomerates of constituent grains with an amorphous silicate core and an organic refractory mantle. Furthermore, H_2O ice was considered to encase the silicate core, organic mantle grains in the outer cold region of the disk ($r \geq 100$ au). They were able to place a tight constraint on the porosity of the agglomerates being approximately $p = 0.95$ or as high as $p = 0.975$. Such highly porous agglomerates have also been successful in fitting the SEDs of dust rings around the A-type star HR 4796A, which is depicted in Fig. 12 and the K-type star ϵ Eridani (Köhler et al. 2008; Li and Lunine 2003; Li et al. 2003). Therefore, it seems plausible that fluffy agglomerates may well represent primitive dust particles not only in the Solar System but also in debris disks.

In a similar approach, Augereau et al. (1999) modeled the HR 4796A debris disk with two distinct populations of dust agglomerates consisting of small grains with a silicate core and an organic mantle. More precisely, agglomerates with a porosity of $p \sim 0.6$ in a cold annulus around 70 au from the star contain amorphous silicate grains and have sizes of $a_c > 10$ μm, while agglomerates with a higher porosity of $p \sim 0.97$ in a warm annulus at about 9 au contain crystalline silicate grains. They used the Maxwell Garnett mixing rule to compute scattering and absorption cross

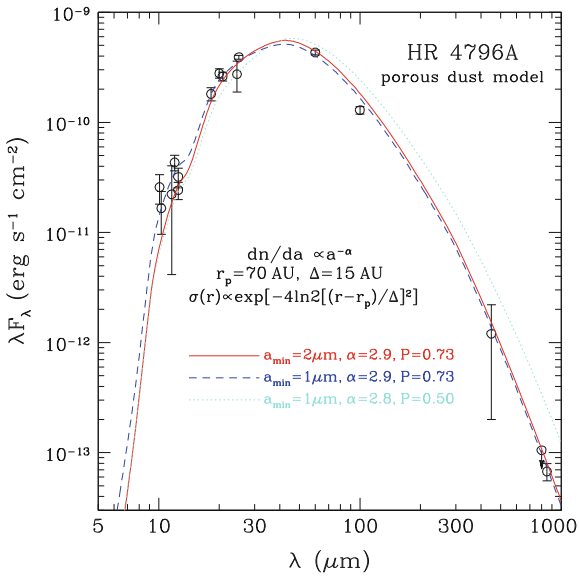


Fig. 12 Comparison of the observed infrared emission of the HR 4796A debris disk to the model spectra calculated from the so-called porous dust model of Li and Greenberg (1998), in which the constituent grains of an agglomerate are composed of amorphous silicate, carbonaceous material, and ice. From Köhler et al. (2008)

sections and the asymmetry parameter of the agglomerates, but the Henyey–Greenstein phase function to simulate the scattered light images. Although the phase function determined by the Maxwell Garnett mixing rule was dismissed, the simultaneous use of the Henyey–Greenstein phase function with the other quantities determined by the Maxwell Garnett mixing rule makes the model not self-consistent.

The young (20 Myr) F-type star HD 181327 was observed by the Herschel Space Observatory in the far-infrared to submillimeter domain providing a detailed coverage of the SED from its cold dust belt (Lebreton et al. 2012, 2014). A simulation of the SED using the Bruggeman mixing rule demonstrated that agglomerates in the dust belt contain $67 \pm 7\,\%$ of ice and are characterized by the porosity $p = 0.63 \pm 0.21$, if the ice is mixed with 2/3 of silicate and 1/3 of amorphous carbon (see Fig. 13). However, Lebreton et al. (2012) recognized that the resulting asymmetry parameter of the agglomerates from the Bruggeman mixing rule is significantly different from the asymmetry parameter determined by a model of scattered light images with the Henyey–Greenstein phase function. While the Henyey–Greenstein phase function does not describe well the phase function of fluffy agglomerates, an estimate of the asymmetry parameter for fluffy agglomerates of submicron constituent grains goes beyond the applicability of the Bruggeman mixing rule.

The same approach was applied to the 30-Myr-old A-type star HD 32297 by Donaldson et al. (2013) who have shown that this system contains fluffy agglomerates with $p = 0.9$. The agglomerates were modeled as a mixture of ices, silicate, and carbon with the volume fraction of 1/2, 1/6, and 1/3, respectively, which

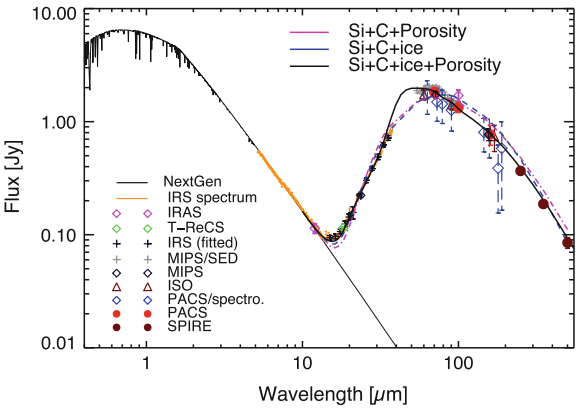


Fig. 13 The spectral energy distribution (SED) of the debris disk around HD 181327 constructed by photometric data (*red crosses* with *error bars*) as well as model SEDs. The best SED fit is achieved by a model of dust agglomerates that are composed of amorphous silicate, carbonaceous material, and ice, while a model without ice or porosity fails to reproduce the observed SED. From Lebreton et al. (2014)

resembles dust in comet 1P/Halley. However, Rodigas et al. (2014) have shown that this cometary dust model fails to reproduce the surface brightness of the disk in the near-infrared wavelength range. The cometary dust model predicts a blue color of the disk in the near-infrared, but the observed brightness of the disk is gray in this wavelength range. It turned out that the near-infrared spectrum of the HD 32297 debris disk is better reproduced by compact ($p = 0$) pure water ice particles, as far as the near-infrared wavelength range is concerned. This is a typical example of how SED fitting models encounter difficulties in providing a warranty for the uniqueness of solutions to the SED fitting problems.

To model the SED of the debris disk around AU Mic from near-infrared to millimeter wavelengths, Fitzgerald et al. (2007) used a Monte Carlo radiative transfer code, which relies on the optical properties of spheres calculated by the Mie theory. Unfortunately, we cannot figure out the reason that such an optically thin debris disk was modeled with the radiative transfer code developed by Pinte et al. (2006) for optically thick protoplanetary disks. According to a model of interstellar dust by Mathis and Whiffen (1989), they assumed porous agglomerates whose constituent grains are composed of silicate, carbon, ices, or "vacuum" and used the refractive indices of the mixture derived from the Bruggeman mixing rule. They also modeled the dependences of intensity and polarization on projected distance from the central star and rejected a model of compact particles. Furthermore, they emphasized the importance of polarimetric data to constrain the composition and distribution of dust particles, but we do not expect that the application of an EMA to agglomerates provides correct understandings for the degree of linear polarization.

5.2 Characteristic Features of Minerals

When high-resolution spectra are obtained in the infrared wavelength range, conspicuous emission feature characteristic of a certain mineral may appear in the spectra. The central wavelength of emission features in the infrared spectra of dust particles is diagnostic of mineral species that are contained in the particles. The most common mineral in primitive dust particles in planetary systems is magnesium-rich olivine, in particular, forsterite, which shows a prominent peak around $\lambda = 11.2\,\mu\text{m}$ (e.g., Campins and Ryan 1989; Knacke et al. 1993). Prominent spectral features of forsterite in primitive dust particles appear not only in the 10 μm wavelength range, but also in the 20–30 μm range (see Fig. 14). A direct link between infrared spectral features and minerals has been confirmed by mineralogical analyses and infrared spectra of CP IDPs (Brunetto et al. 2011).

Hage and Greenberg (1990) used the Maxwell Garnett mixing rule to compute mid-infrared spectra of fluffy agglomerates consisting of cubical-shaped grains, which are randomly distributed in the agglomerates. They considered agglomerates

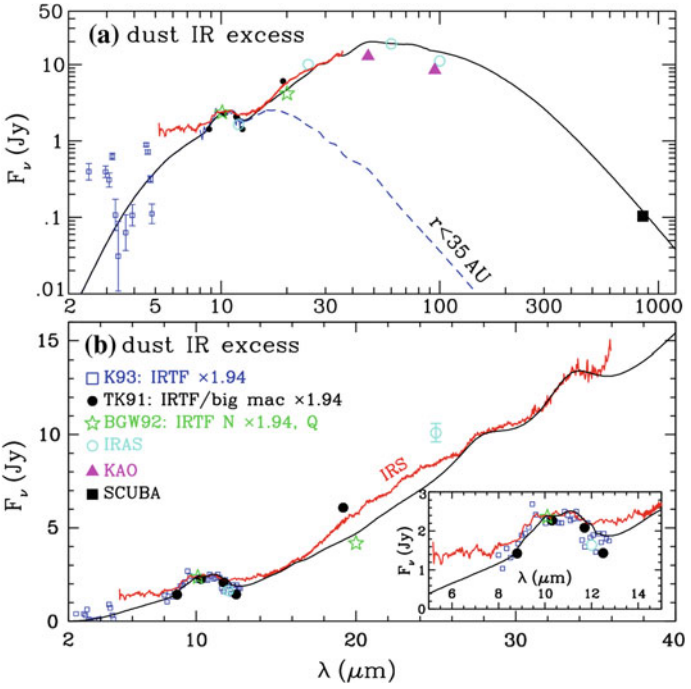


Fig. 14 The spectral energy distribution of the β Pic debris disk and its Spitzer/IRS spectrum (*red line*). In addition to the 10 μm amorphous silicate feature and the 11.2 μm crystalline silicate feature, the 28 and 33.5 μm crystalline silicate features are also prominent in the Spitzer/IRS spectrum. Also shown is the dust infrared emission calculated from the porous dust model of Li and Greenberg (1998). From Chen et al. (2007)

of 1000–8000 cubical identical homogeneous grains in terms of their size, composition, and orientation. In comparison with the DGF/VIEF method of the DDA, they examined the validity of the Maxwell Garnett mixing rule, in which the porosity was defined by the fractional volume of vacuum within the smallest convex volume encasing the agglomerates. They concluded that the porosity of agglomerates in the coma of comet 1P/Halley exceeds 0.97 if agglomerates of $a_v = 10\,\mu m$ contribute to the mid-infrared silicate feature. However, they modeled the 10 μm feature with amorphous silicate, while the silicate emission band observed for comet 1P/Halley indicates that the silicates are crystalline (Bregman et al. 1987). Therefore, their study cannot be used to provide conclusive constraints on the porosity of dust agglomerates in the coma of comet 1P/Halley.

Okamoto et al. (1994) studied the dependences of infrared spectra for fractal agglomerates on the porosity of the agglomerates and the number of their constituent grains. They applied the Maxwell Garnett mixing rule to calculations of absorption cross sections for agglomerates consisting of spherical grains with $a_0 = 0.01\,\mu m$. On the one hand, the infrared spectral features of olivine become weak for large agglomerates with $a_c > 10\,\mu m$ if the agglomerates are relatively compact ($D \sim 3$). On the other hand, the olivine features remain noticeable even for large agglomerates with $a_c > 10\,\mu m$, if the agglomerates are fluffy ($D \sim 2$). This indicates that the size distribution of agglomerates does not influence their infrared spectra, if the agglomerates underwent the growth by the BCCA process because of $D \approx 2$. Therefore, if comets do not exhibit any silicate emission features in their infrared spectra, we may expect that their comae are dominated by large agglomerates of $a_c > 10\,\mu m$ with a relatively compact structure.

Nakamura (1998) attempted to model the characteristic features of olivine in the mid-infrared spectra of the β Pic debris disk using BPCA ($D \sim 3$) particles of $N = 27{,}000$ spherical constituent grains. He assumed the agglomerates to consist of olivine grains and graphite grains half and half with either $a_0 = 0.2\,\mu m$ or $a_0 = 0.5\,\mu m$. The `DDSCAT` code (ver. 4a) with the a_1-term method was used to compute the wavelength dependence of absorption cross section for the agglomerates. The olivine features appear around $\lambda = 10$ and 11 μm for agglomerates with $a_0 = 0.2\,\mu m$ ($a_v = 6\,\mu m$), while the features are weak for those with $a_0 = 0.5\,\mu m$ ($a_v = 15\,\mu m$). However, Golimowski et al. (2006) claim that agglomerates consisting of silicate grains and graphite grains present neutral colors, contrary to red colors observed for the β Pic debris disk. Therefore, there is a room for improvement in modeling of dust particles in the debris disk around β Pic by viewing the observational evidence from various angles.

Kolokolova et al. (2007) considered the evolution of dust mantles on the surface of comets based on the infrared spectra for BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles consisting of spherical grains with $a_0 = 0.1\,\mu m$ computed by the Maxwell Garnett mixing rule. The validity of the Maxwell Garnett mixing rule has been confirmed with the `scsmtm1` code at $N = 1024$ and the `DDSCAT` code (ver. 4a) with the a_1-term method at $N = 32{,}768$. Their results have also shown that olivine features in the infrared spectra of fractal agglomerates disappear at $N > 2^{20}$

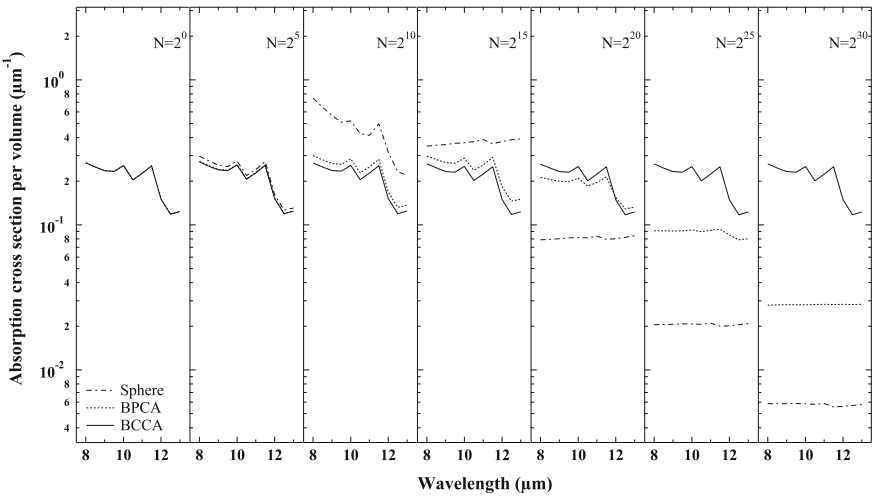


Fig. 15 The dependences of olivine features in the infrared spectra of fractal agglomerates on the number N of constituent grains in an agglomerate and on its structure. The *solid curves*, *dotted ones*, and *dash-dotted ones* are the infrared spectral of BCCA ($D \approx 2$) particles, BPCA ($D \approx 3$) particles, and compact spherical particles, respectively. From Kolokolova et al. (2007)

($a_c > 19$ μm) for BPCA particles of submicron constituent grains (see Fig. 15). The mid-infrared spectral features of olivine are less pronounced for short-period comets with small semi-major axes than those with large semi-major axes (Kolokolova et al. 2007). This implies that short-period comets with small semi-major axes have more compact agglomerates on their surfaces than those with large semi-major axes. This is consistent with the picture of dust mantle formation on the surface of comets, which is characterized by a deficit of highly porous agglomerates on the surface of cometary nuclei due to longer or stronger solar irradiation.

Yamamoto et al. (2008) used fractal agglomerates to interpret the strength of a silicate emission feature as well as the color temperature observed for the ejecta of comet 9P/Tempel 1 during the Deep Impact mission. By taking into account the formation and evolution of dust mantles on comets, they assumed that the dust mantle of the comet consists of agglomerates with $D = 2.5$ and the interior of the comet consists of agglomerates with $D = 1.9$. They used the Maxwell Garnett mixing rule to compute the infrared spectra for the agglomerates whose spherical constituent grains of $a_0 = 0.1$ μm have a structure of a forsterite and amorphous silicate core and an organic refractory mantle. Their success in modeling the infrared spectroscopic observation of the ejecta from comet 9P/Tempel 1 confirmed that the surface layer of periodic comets has been processed and does not maintain their primordial structures and compositions (cf. Kolokolova et al. 2007).

It is well known that the position of any emission feature characteristic of a certain mineral in general depends on the shape of the mineral grains (e.g., Bohren and Huffman 1983, § 12.2.7). Yanamandra-Fisher and Hanner (1999) employed the

DDSCAT code (ver. 4a) with the LDR method to study the effect of grain shapes on the positions of silicate features in the infrared spectra. They considered small agglomerates of 2–5 submicron silicate constituent grains whose shapes are either spherical or tetrahedral. Their results show strong shape effects on the positions of forsterite features as expected, but the shape effects are weak for infrared spectral features of amorphous silicate. Since absorption bands take place at wavelengths where the real part of complex dielectric function is negative, shape effects are not important if the negative value is confined in a very narrow wavelength range. Infrared spectral observations of comets and debris disks have shown that the positions of forsterite features appear at similar wavelengths, irrespective of their different circumstances (e.g., Knacke et al. 1993). Therefore, the complex dielectric function of primitive dust particles in planetary systems most likely does not have a deep negative value compared to pure forsterite, indicating that the particles in planetary systems are not composed of bare mineral grains.

The presence of organic refractory material encasing mineral grains plays a vital role in the positions of the peaks that appear in the infrared spectra (Kimura 2013). This is known as the matrix effect, which has been experimentally proven to shift the positions of olivine features to longer wavelengths (Day 1975; Dorschner et al. 1978). Kimura (2013) computed the infrared spectra of BCCA ($D \approx 2$) particles consisting of silicate core, organic mantle spherical grains with $a_0 = 0.1$ μm using the Maxwell Garnett mixing rule. It turned out that the positions of olivine features are consistent with the infrared spectra of dust in cometary comae, debris disks, and protoplanetary disks, if the organic volume fraction fulfills the cosmic abundance constraints (see Fig. 16). This indicates that the composition of agglomerates is not

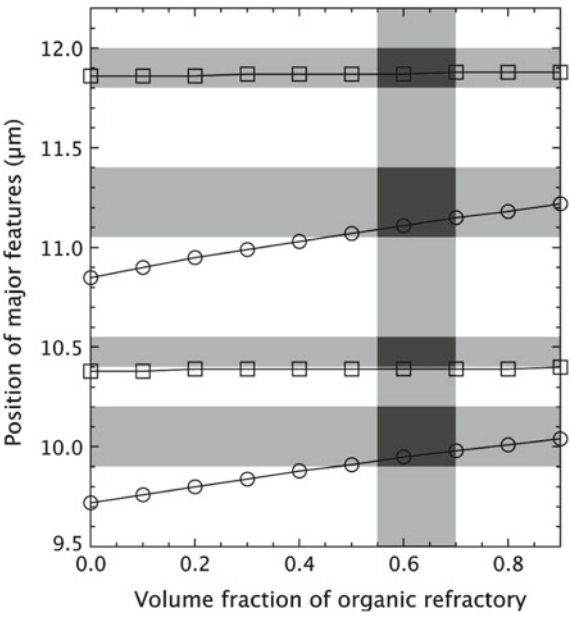


Fig. 16 The central wavelengths of forsterite features in simulated infrared spectra of BCCA ($D \approx 2$) particles versus the volume fraction of an organic refractory mantle that encases a forsterite core in spherical constituent grains. The *horizontally long shaded area* indicates the central wavelengths of forsterite features in observed infrared spectra of primitive dust particles in planetary systems. The *vertically long shaded bar* is the plausible range of volumetric organic fraction inferred from the cosmic abundance constraints. From Kimura (2013)

completely a free parameter in modeling primitive dust particles in planetary systems.

It is worth noting that organic refractory material could be carbonized due to the preferential loss of hydrogen, nitrogen, and oxygen by ultraviolet irradiation and ion bombardments (Jenniskens 1993; Jenniskens et al. 1993). Since the carbonization of organic refractory material changes its chemical composition, the refractive indices of organic-rich carbonaceous material depend on the degree of carbonization. Therefore, the degree of carbonization affects the positions of mineral features in the infrared spectra, if the mineral is encased in organic-rich carbonaceous material. Using the Maxwell Garnett mixing rule, Kimura (2014) has studied the effect of carbonization on the infrared spectral features of olivine for BCCA ($D \approx 2$) particles of $a_0 = 0.1\,\mu\text{m}$ spherical constituent grains with a forsterite core and an organic-rich carbonaceous mantle (see Fig. 17). The forsterite feature at $\lambda = 11.1\,\mu\text{m}$ observed in the debris disk of β Pic indicates that dust particles in the β Pic disk did not suffer from severe carbonization, compared to cometary dust in the Solar System. Therefore, infrared spectra of dust particles with high spectral resolution are highly useful to diagnose the carbonization degree of organic materials in the particles.

Infrared spectroscopic observations of comets reported on the discovery of a 9.3 μm feature in the coma of comets C/1995 O1 (Hale–Bopp), C/2001 Q4 (NEAT), and 9P/Tempel 1 (e.g., Wooden et al. 1999, 2004; Harker et al. 2007). Although the 9.3 μm feature is not always present in the spectra, the feature has been attributed to magnesium-rich pyroxene. Kimura et al. (2008) demonstrated that the 9.3 μm feature could also be produced by forsterite, if amorphous silicate constituent grains are covered by a forsterite layer encased in organic refractory mantle (see Fig. 18). While the presence of pyroxene in primitive dust particles could be supported by mineralogical investigation of CP IDPs, one should keep in mind that the mineralogical identification with an infrared feature alone is not always unique.

6 Integral Optical Quantities of Dust Agglomerates

6.1 Bolometric Albedo

The bolometric albedo of dust particles is an integrated quantity over wavelengths, defined as the fraction of stellar luminosity scattered by the particles (Gehrz and Ney 1992). Therefore, the SED could be used to estimate the bolometric albedo of dust particles, when observational data on scattered light are available. The bolometric albedos of dust particles in comet 1P/Halley and the debris disks of β Pictoris, HD 207129, and HD 92945 have been determined to be 0.32, 0.35, 0.051, and 0.10, respectively (Gehrz and Ney 1992; Backman et al. 1992; Krist et al. 2010; Golimowski et al. 2011).

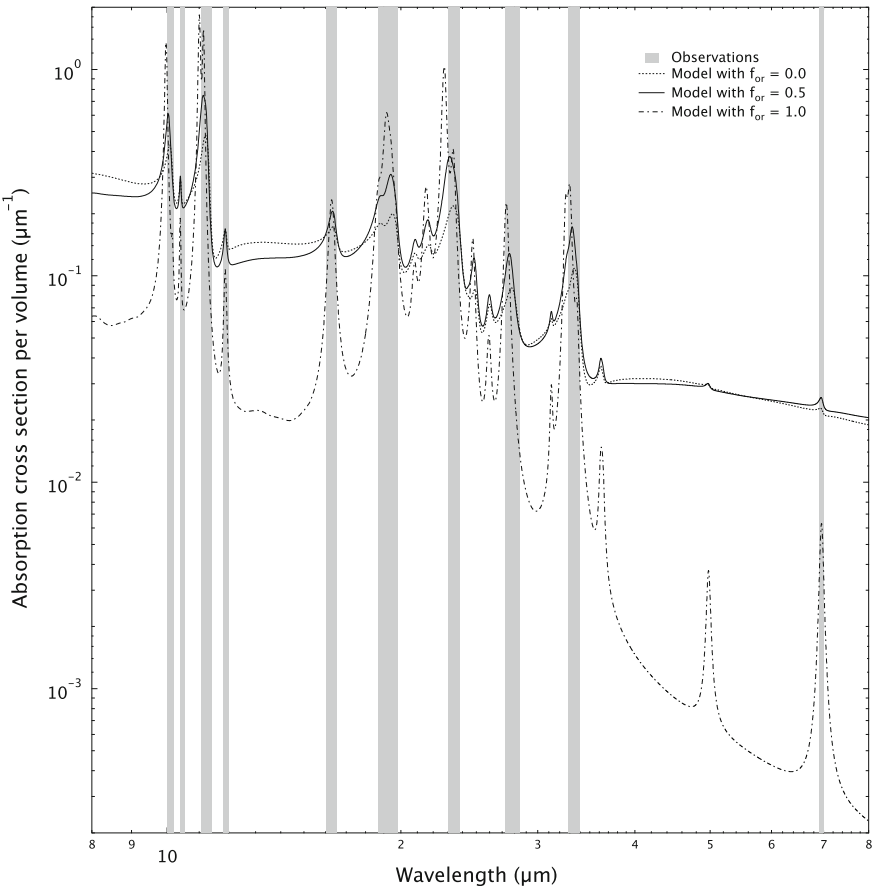


Fig. 17 Infrared spectra of dust agglomerates consisting of 2^{30} submicron grains with a forsterite core and an organic refractory mantle. The carbonization of organic refractory mantle is characterized by $1 - f_{or}$ where f_{or} denotes the volume fraction of organic refractory material within the carbonaceous mantle. The *shaded bars* indicate the central wavelengths that have been observed to exhibit noticeable emission features from dust particles in comets, debris disks, and protoplanetary disks. From Kimura (2014)

It is sometimes claimed that numerical simulation of the SED with an EMA predicts a much higher albedo than the bolometric albedo from the SED (Krist et al. 2010; Lebreton et al. 2012). This discrepancy has been attributed to the inapplicability of the EMA to properly describe the light-scattering properties of agglomerates in planetary systems. We have no objection to the note of caution that EMAs are not applicable to agglomerates of submicrometer-sized constituent grains in the visible wavelength range. However, we notice that the simulated albedo is not the bolometric albedo, but the single-scattering albedo, which is defined by the ratio of the scattering cross section to the extinction cross section. Therefore, to the best

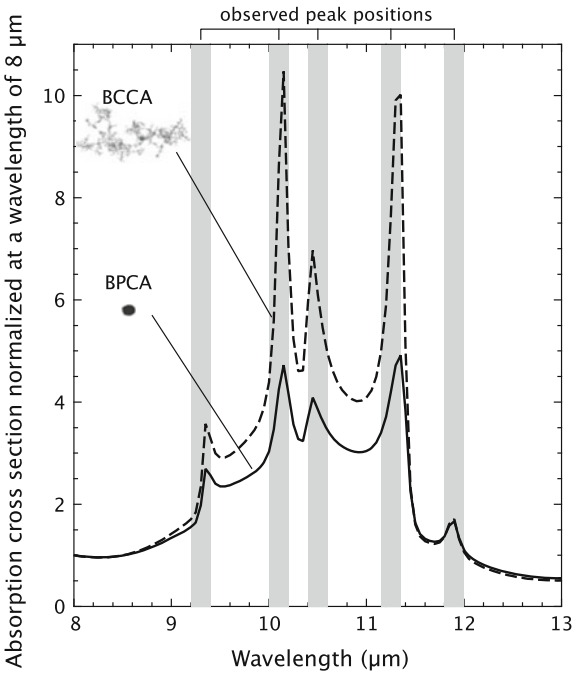


Fig. 18 Infrared spectra of dust agglomerates consisting of 2^{20} submicron grains with an amorphous silicate core and an organic refractory mantle after processing by exothermic chemical reactions. The processing of such a core–mantle grain forms a thin forsterite layer on the surface of amorphous silicate cores. The *shaded bars* indicate the central wavelengths that have been observed to exhibit noticeable emission features from dust particles in cometary comae. From Kimura et al. (2008)

of our knowledge, there has not yet been a direct comparison of the bolometric albedo of primitive dust particles in planetary systems between numerical simulations and astronomical observations.

6.2 Radiation Pressure

Stellar radiation inevitably exerts a force on a dust particle in planetary systems and thus an estimate of radiation pressure has been of great importance to understand the dynamics of dust particles in planetary systems. An estimate of radiation pressure requires integration of radiation pressure cross sections weighted by stellar radiation spectrum over wavelengths from ultraviolet to far-infrared. Stellar radiation pressure also retards the motion of dust particles and, in consequence, the particles spiral to the central star, referred to as the Poynting–Robertson effect (Robertson 1937). In particular, stellar radiation pressure controls the dynamics of dust particles and their spatial distribution, unless the stellar mass loss rate is more than a few times larger than the solar mass loss rate or the total mass of a circumstellar dust disk is heavier than one millionth of the Earth (Minato et al. 2006).

It is a common practice to discuss the importance of radiation pressure in terms of β defined by the ratio of stellar radiation pressure to gravitational attraction. A pioneering work on numerically evaluating the β values for fluffy agglomerates

was performed by Mukai et al. (1992). They applied the Maxwell Garnett mixing rule to estimate effective refractive indices for fractal agglomerates and then computed radiation pressure cross sections in the framework of Mie theory. Mukai et al. (1992) assumed $a_0 = 0.01\,\mu\text{m}$, because the Maxwell Garnett mixing rule is applicable under the condition that the radius of inclusions is much smaller than the wavelength of light. It turned out that the β values computed by the EMA are in good agreement with those computed by the DDSCAT code (ver. 4a) with the DGF/VIEF method for $N \leq 1024$, except for highly porous ($D \approx 2$) agglomerates composed of silicate constituent grains (Mukai et al. 1992). Kimura et al. (1997) applied the Bruggeman mixing rule to compute the β values for BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles of 0.01 μm-radius constituent grains. Their results are also shown to agree with the results computed by the DDSCAT code (ver. 4a) with the a_1-term method, except for silicate BCCA particles (Kimura et al. 2002a). Therefore, EMAs provide reasonably accurate results for β values of agglomerates as far as a relatively compact agglomerate or any agglomerate composed of absorbing material is concerned.

In general, the dependence of β values on the size of agglomerates becomes weaker as the fractal dimension of the agglomerates decreases (Mukai et al. 1992; Kimura et al. 1997, 2002a). Therefore, β values of agglomerates approach those of their constituent grains with increase in the porosity of the agglomerates, even if the agglomerates are not fractals. This statement agrees with the results of Saija et al. (2003) who applied the TMM to compute the β values for non-fractal agglomerates. They considered 200 constituent grains of silicate or amorphous carbon with $a_0 = 0.005\,\mu\text{m}$ and the cluster of the grains with $p = 0.842-0.995$ in circumstellar environments with stellar temperatures of 2700, 5800, and 10,000 K. As a result, the β values of agglomerates are smaller than those of volume equivalent compact spheres in the submicron-size range and larger in the sizes above tens of microns.

Wilck and Mann (1996) used the Maxwell Garnett mixing rule for their computations of the β values for porous spheres that mimic fluffy agglomerates in the Solar System. A similar approach to the computation of β values using the Maxwell Garnett mixing rule has been adopted for porous spheres around A-type stars (Artymowicz and Clampin 1997; Grigorieva et al. 2007). Kirchschlager and Wolf (2013) used the DDSCAT code (ver. 7.1) with the LDR method to compute the β values for porous silicate spheres with a fixed porosity in the environments of the Sun and other main-sequence stars. Kimura and Mann (1999a) have shown that the β values for small ($N \leq 2048$) BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles of amorphous carbon and silicate computed by the DDSCAT code (ver. 4a) with the a_1-term method could be well represented by those of porous spheres with a fixed porosity computed by the Bruggeman mixing rule as far as small constituent grains of $a_0 = 0.01\,\mu\text{m}$ are concerned. It is worth noting that the β values of porous spheres with a fixed porosity are inversely proportional to the radius of the spheres in the sizes above tens of microns, independent of the porosity. Such a size dependence of porous spheres differs from that of fractal agglomerates, in particular, highly porous ones ($D \approx 2$) whose β values tend to be nearly independent of their sizes.

The major drawback of the above-mentioned studies is the assumption that the radius of constituent grains is smaller than the wavelength of ultraviolet light, owing to the limitation for the EMAs, the DDA a_1-term method, or computational resources such as RAM and CPU time. However, the constituent grains of agglomerates in planetary systems are one order of magnitude larger than $a_0 = 0.01$ μm, as pointed out in Sect. 2. A deficit of small dust particles in planetary systems takes place around sizes near the maximum β values due to mutual collisions of the particles and subsequent blowouts by stellar radiation pressure (Ishimoto and Mann 1999; Krivov et al. 2000). In conclusion, correct estimates of the β values for agglomerates of submicron constituent grains are essential to better understand not only the dynamical behavior but also the spatial distribution of dust particles in a planetary system.

Kimura et al. (2002b) computed the β values for agglomerates of submicron silicate constituent grains using the `scsmtm1` code in order to simulate dust particles released from sungrazing comets. Although the number of constituent grains was limited to $N \leq 32$ by computing resources for the `scsmtm1` code, their results show that agglomerates of $a_0 \approx 0.1$ μm are consistent with the dynamical constraint of $\beta \leq 0.6$ for dust particles in the tails of mini ($a \sim 1-10$ m) sungrazers (cf. Sekanina 2000). However, this does not mean that dust particles in sungrazing comets are too peculiar to lack organic refractory materials. The assumption of bare silicate grains is easily justified, since the organic refractory component sublimates almost instantly in the vicinity of the Sun. In the tail of the larger ($a \sim 100$ m) sungrazing comet C/2011 W3 (Lovejoy), the values of $2.5 > \beta \geq 0.6$ were identified before perihelion, although only $\beta \approx 0.6$ appeared after perihelion (Sekanina and Chodas 2012). Therefore, the high β values can be associated with organic-rich carbonaceous materials that are less refractory than silicates with $\beta \approx 0.6$. We cannot help wondering whether sublimation of organic materials is also responsible for unusual degrees of linear polarization observed for sungrazing comets (cf. Weinberg and Beesonn 1976; Thompson 2015).

Mukai and Okada (2007) studied the size dependence of the β values for BPCA ($D \approx 3$) particles of silicate constituent grains with $N = 2048$ and 16,384, by varying the radius of constituent grains. They used the Maxwell Garnett mixing rule to compute the radiation pressure cross sections of the agglomerates, although the radius of constituent grains ranged from nanometer to millimeter. Because the size dependence of the β values in their results originates from the variation in the radius of constituent grains, the β values were independent of N. By extending the study of Mukai and Okada (2007) to agglomerates of amorphous carbon constituent grains, Levasseur-Regourd et al. (2007) concluded that the β values depend on the composition of the grains even for millimeter-sized agglomerates. Although their results would have important implications for the dynamics of millimeter-sized dust particles in cometary trails, we should point out that the size dependence of the β values with a fixed number of constituent grains differs from that with a fixed radius of constituent grains.

Köhler et al. (2007) extended β values of agglomerates with $a_0 = 0.1$ μm up to $N = 512$ for BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles of silicate constituent

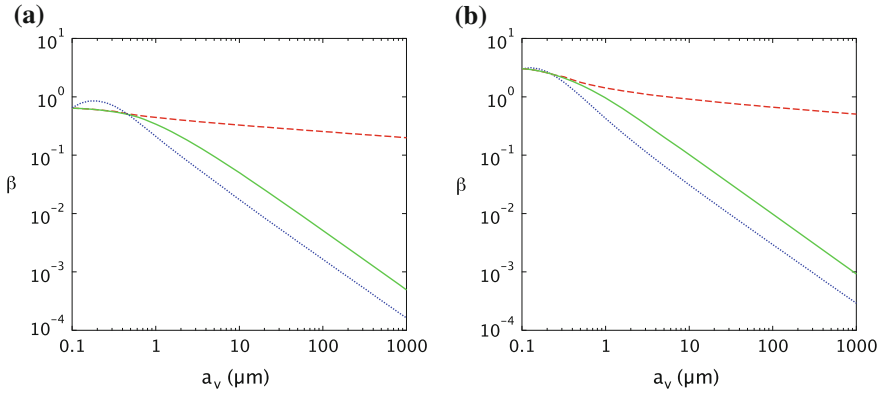


Fig. 19 The ratios of radiation pressure to gravitational attraction, β, of the Sun acting on **a** silicate particles and **b** carbon particles as a function of a_v, the radius of volume equivalent spheres. *Solid line*: fractal agglomerates consisting of 0.1 μm-radius particles grown by BPCA ($D \approx 3$); *dashed line*: fractal agglomerates consisting of 0.1 μm-radius particles grown by BCCA ($D \approx 2$); *dotted line*: compact spheres. Reproduced from Fig. 5 of Minato et al. (2006)

grains or carbon constituent grains. They used both the DDSCAT code (ver. 6.1) with the LDR method and the scsmtm1 code for the computations of radiation pressure cross sections. The radiation pressure cross sections were computed in the range of wavelengths from ultraviolet to far-infrared for small agglomerates of $N \leq 32$, but only at a wavelength $\lambda = 0.6\,\mu\text{m}$ for larger agglomerates of $32 < N \leq 512$. The approximation to the β values at a single wavelength of 0.6 μm was shown to reproduce those computed with a full wavelength range within the accuracy of 30 %. Furthermore, Köhler et al. (2007) extrapolated the results to much larger ($N \leq 10^{17}$) agglomerates on the assumption that the radiation pressure cross sections for large agglomerates are proportional to $N^{2/D}$ in the geometrical optics regime. The overall size dependence of the β values for agglomerates of submicron grains has significant implications for a better understanding of their dynamical behaviors in planetary systems.

It has been known that the geometric cross sections of fractal agglomerates show a deviation from the ideal proportionality to $N^{2/D}$ (Meakin and Donn 1988; Ossenkopf 1993). Therefore, Minato et al. (2006) took into account the geometric cross sections of the agglomerates to improve the results of Köhler et al. (2007) for $N > 512$. In Fig. 19, we reproduce their results[2] on the overall size dependence of β values for agglomerates of submicron silicate constituent grains (left) and those of submicron carbon constituent grains (right). Their results were also applied to estimate the ratio of solar radiation pressure to the local gravity of sub-km sized

[2]The size dependence of the β values for agglomerates of submicron constituent grains has not explicitly been shown in Minato et al. (2006), but could be deduced from their figures.

asteroids acting on BPCA and BCCA particles and their β values in the debris disk of the A-type star Fomalhaut (Kimura et al. 2014). Stellar radiation pressure could replenish BPCA particles of $a_{\rm v} = 10-370\,\mu$m into the exo-zodiacal cloud of Fomalhaut, if their parent bodies orbit the star at 2 au. Accordingly, the size dependence of β values for agglomerates of submicrometer-sized constituent grains play a vital role in better understanding of dust dynamics in planetary systems.

Tazaki and Nomura (2015) applied the `scsmfo1b` code with the QMC method to compute the β values for agglomerates of spherical constituent grains in orbit around a star with a blackbody radiation of 5778 K. Note that the solar radiation spectrum is close to a spectrum of a 5778 K blackbody and thus their results can be applied to dust particles in the Solar System. The agglomerates were assumed to consist of amorphous silicate spheres with a radius of $a_0 = 0.01$ or 0.1 μm and have grown under the BCCA ($D \approx 2$) process in the range of $N \leq 1024$. Their results coincide with those in Kimura and Mann (1999a) and Minato et al. (2006), which were obtained by the `DDSCAT` code (ver. 4a) with the a_1-term method and the `DDSCAT` code (ver. 6.1) with the LDR methods, respectively. This coincidence is a natural consequence, because the computational techniques are not critical to the integrated optical quantities such as β.

Kimura et al. (2003b) applied the `gmm02TrA` code to compute the β values for BCCA ($D \approx 2$) particles consisting of submicron silicate core, organic mantle grains in the Solar System. They have shown that the β values for the agglomerates of submicron silicate core, organic mantle grains fulfill the dynamical constraint on the β values for interstellar dust streaming into the Solar System (see Fig. 20). However, it turned out that the agreement happened by coincidence, since interstellar dust in the Solar System does not contain organic material (Westphal et al. 2014; Altobelli et al. 2016). Nevertheless, the results of Kimura et al. (2003b) are applicable to cometary dust and dust in debris disks, although their results are available only up to $N = 16$. Indeed, the β values for the agglomerates of submicron silicate core, organic mantle grains are consistent with $\beta < 2.5$ derived from coronagraphic images of comet C/2011 W3 (Lovejoy) by Sekanina and Chodas (2012).

Silsbee and Draine (2016) computed the β values for silicate agglomerates with $N = 32$ and 256 using a blackbody radiation of 5780 K. Although they aimed to study the dynamics of interstellar dust streaming into the Solar System, their computed β values can be applied to dust particles in planetary systems. They used the `MSTM` code (ver. 3.0) and the `DDSCAT` code (ver. 7.3) with the "modefied" LDR method for BPCA particles with and without restructuring. The radius of the agglomerates was assumed to lie in the range of $a_{\rm v} = 0.01-1\,\mu$m for $N = 32$ and $a_{\rm v} = 0.1-0.8\,\mu$m for $N = 256$, implying $a_0 = 0.0032-0.32\,\mu$m and $a_0 = 0.016-0.126\,\mu$m, respectively. Their results show that β values for silicate agglomerates do not exceed unity, in concord with previous studies. They have also shown that the substitution of iron constituent grains for some of silicate constituent grains in an agglomerate enhances β values. According to their results, the condition of $\beta > 1$ requires more than 35 % of the volume be iron, but this condition conflicts with the cosmic abundances of iron, silicon, and magnesium.

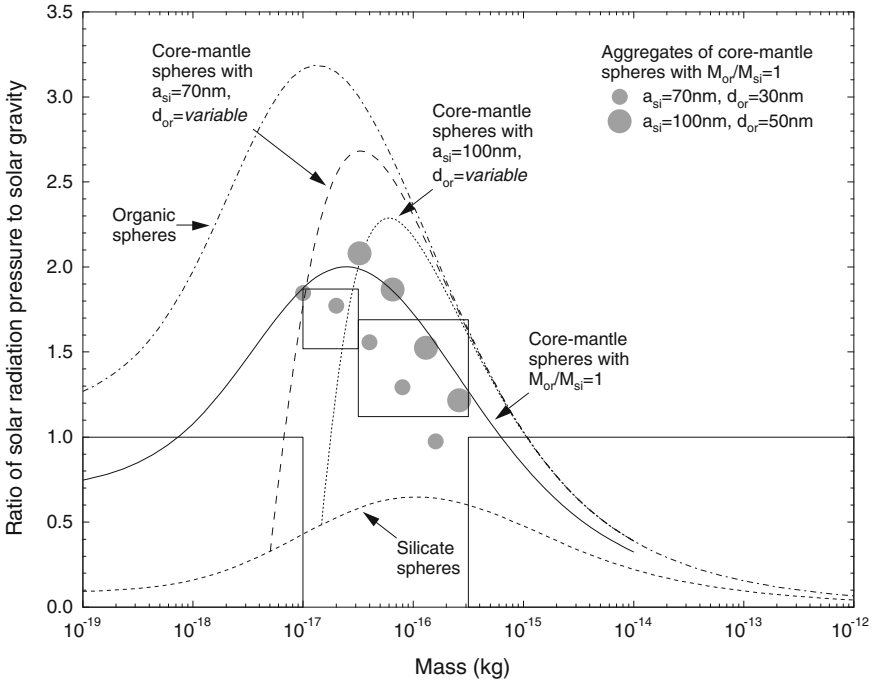


Fig. 20 The ratios of radiation pressure to gravitational attraction, β, of the Sun acting on BCCA ($D \approx 2$) particles consisting of submicron core–mantle particles. The core and the mantle of the constituent spherical grains are composed of amorphous silicate and organic refractory material, respectively, in equal mass. Also plotted are the β ratios for bare amorphous silicate spheres, bare organic refractory spheres, and silicate core, organic mantle spheres. The *rectangles* indicate the expected ranges of β values derived from in situ impact data of interstellar dust by Ulysses. From Kimura et al. (2003b)

Stellar radiation exerts a force on a sphere in the direction antiparallel to the stellar gravity, but dust agglomerates are also subjected to a component of force perpendicular to the propagation of stellar radiation (e.g., van de Hulst 1957, § 2.3). Kimura and Mann (1998) applied the DDSCAT code (ver. 4a) with the a_1-term method to compute the radiation pressure cross sections for BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles of 0.01 μm-radius constituent grains in the directions parallel and perpendicular to the propagation of stellar radiation. Kimura et al. (2002a) determined the parallel and perpendicular components of the β values for BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles of 0.01 μm-radius constituent grains in the Solar System using the DDSCAT code (ver. 4a) with the a_1-term method (see Fig. 21). The DDSCAT code (ver. 4a) with the a_1-term method was also used to compute the parallel and perpendicular components of radiation pressure on BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles of 0.01 μm-radius constituent grains in orbit around β Pic (Kimura and Mann 1999b). The perpendicular components of radiation pressure are negligible for carbon agglomerates in comparison to the parallel

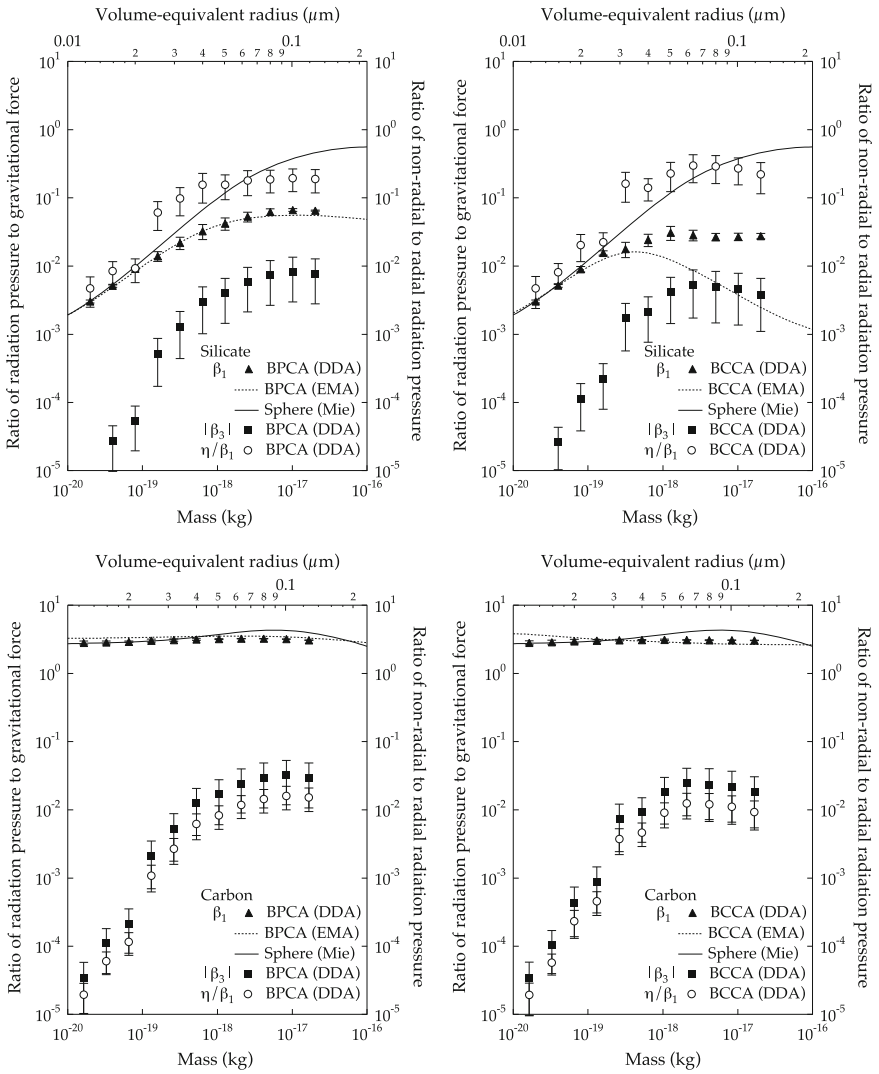


Fig. 21 The ratios of radiation pressure to gravitational attraction of the Sun acting on fractal agglomerates consisting of 0.01 μm-radius grains. The *filled triangles* and *squares* are the ratios in the direction antiparallel to the solar gravity and the direction perpendicular to the orbital plane of the agglomerates. The *open circles* indicate the transverse component divided by the radial component of solar radiation pressure. The *dashed curves* are numerical results on the radial component of solar radiation pressure relative to the solar gravity estimated by the Bruggeman mixing rule. The *solid curves* are the radial components of solar radiation pressure relative to the solar gravity acting on compact spherical particles. From Kimura et al. (2002a)

component, but could be significant for silicate agglomerates. However, we are not very confident that this conclusion holds for agglomerates consisting of submicron grains prior to a future study with submicron constituent grains.

Silsbee and Draine (2016) applied the DDSCAT code (ver. 7.3) with the "modefied" LDR method to compute the transverse component of radiation pressure on silicate BPCA particles with restructuring in the Solar System. They assumed that the agglomerates with $N = 32$ have a radius $a_v = 0.15$ and 0.6 μm, implying $a_0 = 0.047$ and 0.19 μm, respectively. Their results would suggest that the transverse component gives a more effective consequence for the dynamics of smaller agglomerates. However, we cannot assert at this stage that this is a general trend, because it is not clear whether the difference originates from smaller a_v or smaller a_0.

6.3 Equilibrium Temperature

Dust particles in planetary systems attain equilibrium temperatures depending on their composition, structure, and distance from the central star. Equilibrium temperature is a key parameter to determine the SEDs as well as the sublimation zone of dust particles in the vicinity of a star. In equilibrium, the temperature of dust particles in planetary systems is determined by the balance of energies between absorption of stellar radiation and thermal emission and sublimation of the particles. An estimate of equilibrium temperature requires integration of absorption cross sections weighted by stellar radiation spectrum and the Planck function over wavelengths from ultraviolet to far-infrared. It is, therefore, not an easy task to estimate the equilibrium temperature of agglomerates consisting of submicrometer-sized grains in particular.

Kozasa et al. (1992) used an empirical formula to calculate the equilibrium temperature of BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles consisting of magnetite grains with a radius $a_0 = 0.01$ μm at 1 au from the Sun. They proposed a formula to approximate the absorption cross sections for agglomerates using the optical characteristics of a spherical cloud that encloses the constituent grains. This formula reproduces the results from the DDSCAT code (ver. 4a) with the DGF/VIEF method within about 10 %, if the agglomerates are assumed to have the same radius as a sphere of equal geometric cross section. It should be, however, noted that the empirical formula is limited to agglomerates of tiny constituent grains with $a_0 \leq 0.01$ μm.

Mann et al. (1994) estimated the equilibrium temperature of BPCA ($D \approx 3$) particles consisting of silicate grains and graphite grains with a radius $a_0 = 0.01$ μm near the Sun using the Maxwell Garnett mixing rule. A higher temperature was found for composite agglomerates of silicate grains and graphite grains in comparison to pure silicate agglomerates or pure graphite agglomerates. The high temperature of an agglomerate consisting of silicate grains and carbon grains was confirmed by Kimura et al. (1997) who applied the Bruggeman mixing rule,

although the formula for energy of sublimation in Kimura et al. (1997) is different from that in Mann et al. (1994). Whereas Mann et al. (1994) assumed that the contribution of each constituent grain to sublimation is given by its volume fraction to the power of 2/3, Kimura et al. (1997) describe that the contribution of each constituent grain to sublimation is proportional to its volume fraction. We should point out that the incorrectness of the formula in Mann et al. (1994) is obvious, as it fails to work if two materials have an even volume fraction.

Kimura et al. (1997) found that the equilibrium temperatures of fractal agglomerates with $a_0 = 0.01$ μm become less dependent on their sizes as the fractal dimension decreases or the porosity increases. As a result, the temperatures of fractal agglomerates asymptotically approach that of constituent grains in the limit of $D \to 2$. This trend is associated with the fact that both the absorption cross section and surface area of fractal agglomerates are proportional to the number of constituent grains if $D = 2$. Therefore, we expect for fractal agglomerates that the size dependence of their temperatures is weak and close to that for the constituent grains, despite the radius of the grains, as long as agglomerates of $D \approx 2$ are concerned.

Again, the assumption of $a_0 = 0.01$ μm for the radius of constituent grains is the major drawback of the above-mentioned studies as a simulation for primitive dust particles in planetary systems. To our knowledge, Greenberg and Hage (1990) performed the earliest study on the equilibrium temperature of fluffy agglomerates consisting of submicron grains in an application to dust particles in comet 1P/Halley at 0.9 au from the Sun. They assumed that the constituent grains of $a_0 = 0.1$ μm have a structure of a silicate core and an organic mantle and used the Maxwell Garnett mixing rule for wavelengths of $\lambda \geq 1$ μm and, for shorter wavelengths, the optical characteristics of a spherical cloud that encloses the constituent grains. They claimed the assumption of the cloud to be a good approximation as long as the size parameter of the cloud radius exceeds 10, but did not show any evidence for their claim. The dependences of temperature on porosity and agglomerate size are qualitatively the same as, but quantitatively different from the case for $a_0 = 0.01$ μm. Therefore, the use of submicrometer-sized constituent grains in an agglomerate for computation of equilibrium temperature is of crucial importance.

Greenberg and Li (1998) calculated the equilibrium temperature of fluffy agglomerates consisting of spherical silicate core, organic refractory mantle grains with $a_0 = 0.1$ μm at 1 au from the Sun. They presented how the temperature depends on the mass and porosity of the agglomerates and on the mass ratio of the silicate core and the organic refractory mantle. Their calculations were performed with the Maxwell Garnett mixing rule, but we notice that the Maxwell Garnett mixing rule alone is not applicable to calculations of temperatures for $a_0 = 0.1$ μm, because the calculations require the absorption cross sections at wavelengths of $\lambda < 1$ μm where $x_0 \geq 1$.

Li and Greenberg (1998) estimated the radial distribution of equilibrium temperatures for agglomerates of submicron constituent grains in β Pic (see Fig. 22).

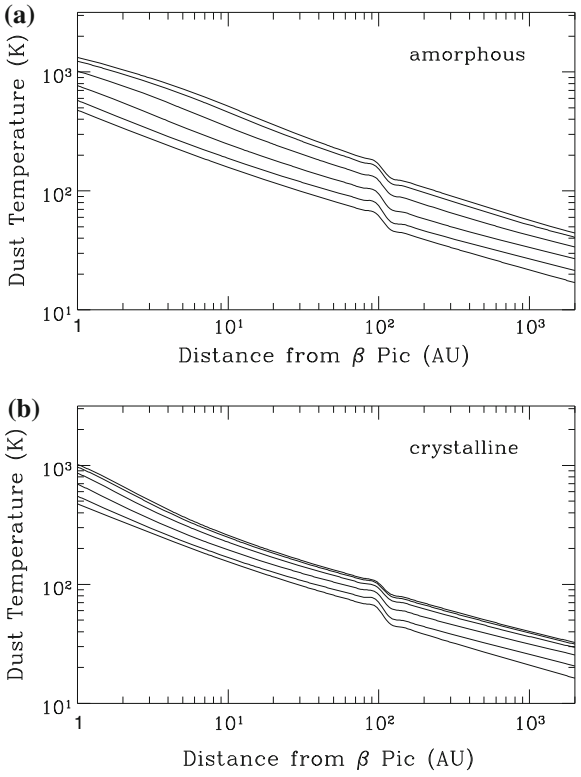


Fig. 22 The equilibrium temperature of porous agglomerates with masses of $m = 10^{-17}$, 10^{-15}, 10^{-13}, 10^{-10}, 10^{-7} kg (from *top* to *bottom*) as a function of distance from the central star in the debris disk of β Pic. Spherical constituent grains of an agglomerate have a radius $a_0 = 0.1$ μm and comprise either **a** amorphous silicate encased in an organic refractory material (*upper panel*) or **b** crystalline silicate (*lower panel*). In the outer region of the disk at distances $r > 100$ au, the agglomerates are assumed to have an external layer of H_2O ice. From Li and Greenberg (1998)

They assumed that the constituent grains with $a_0 = 0.1$ μm are composed of either an amorphous silicate core and an organic mantle or bare crystalline silicate. The presence of ice mantles is also considered if the agglomerates are located at a distance exceeding 100 au from the star, implying ice sublimation at 100 au. According to the method described in Greenberg and Hage (1990), they used the Maxwell Garnett mixing rule for $\lambda \geq 1$ μm and considered a cloud of the constituent grains for $\lambda < 1$ μm. Their results indicate that sublimation of ices in the inner disk enhances the temperatures of the agglomerates as a result of the increase in absorptivity.

Kimura et al. (2002b) computed the equilibrium temperatures of BCCA ($D \approx 2$) particles consisting of submicron silicate grains in the vicinity of the Sun using the `scsmtm1` code. Although their computations with $a_0 = 0.1$ μm were limited to $N \leq 32$, their results confirmed that the temperatures of agglomerates show a weaker size dependence, compared to those of compact spheres (see Fig. 23). They have shown that the temperatures of pyroxene agglomerates are 0.5–0.6 times lower than those of olivine agglomerates in the comae and tails of sungrazing comets. Accordingly, their results predict that crystalline and amorphous pyroxene grains sublimate inside five solar radii ($R_\odot$) from the Sun, while crystalline and amorphous

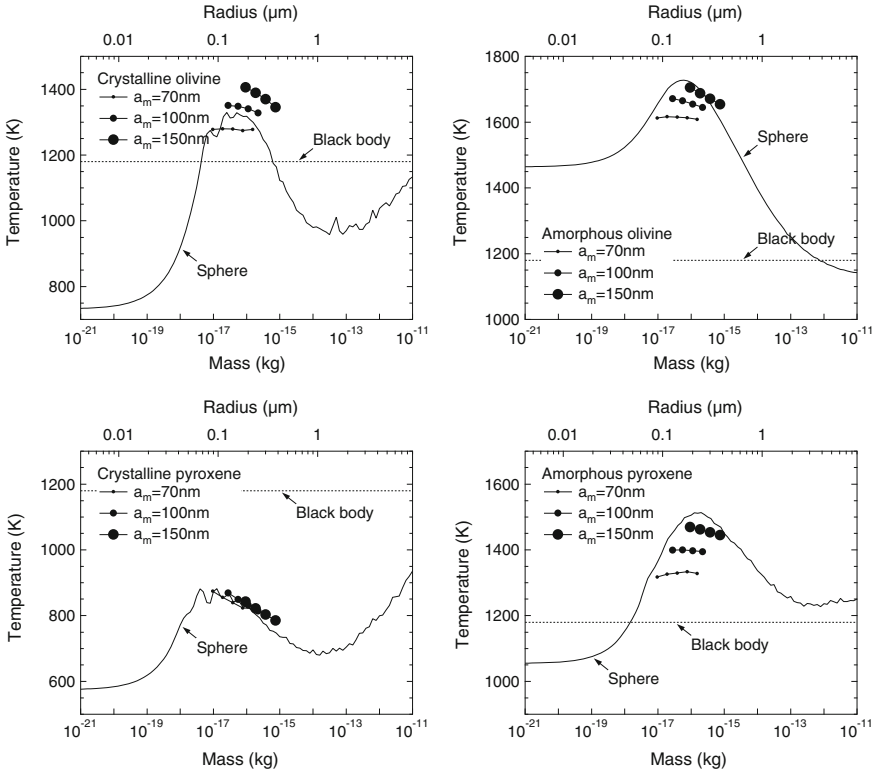


Fig. 23 The equilibrium temperature of fractal agglomerates consisting of spherical olivine and pyroxene grains at a heliocentric distance of $12\,R_\odot$. The *smallest*, the *medium*, and the *largest filled circles* are the numerical values of the temperature with $a_0 = 0.07,\ 0.10,$ and $0.15\,\mu$m, respectively. The *solid curves* and *dotted ones* are the equilibrium temperature of compact spherical particles and the blackbody temperature. From Kimura et al. (2002b)

olivine grains sublimate inside $10 - 11 R_\odot$ and $12 - 13 R_\odot$, respectively. It turned out that sublimation of pyroxene agglomerates and olivine ones near the Sun well accounts for not only light curves of sungrazing comets observed by SOHO/LASCO but also neutral hydrogen Lyman α emission from sungrazers observed by SOHO/UVCS (cf. Biesecker et al. 2002; Bemporad et al. 2006, 2007).

Xing and Hanner (1996), (1997) estimated the dependence of equilibrium temperatures on heliocentric distance for agglomerates consisting of 10 submicron spherical carbon grains or tetrahedral ones with various separations of the grains. They used the DDSCAT code (ver. 4b) with the LDR method for computing absorption cross sections of the agglomerates in the wavelength range from visible to far-infrared. The absorption cross sections for ultraviolet and visible wavelengths were extrapolated from the cross section at $\lambda = 1\,\mu$m on the assumption that the absorption cross section at $\lambda = 0$ equals to the geometric cross section of the

agglomerates. Their results indicate that the temperature of agglomerates approaches that of constituent grains once the grains get separated, while the results for touching grains are similar to those for overlapping grains. Agglomerates consisting of tetrahedral grains show a weaker dependence of temperature on the separation than agglomerates consisting of spherical grains. In addition, the dependence of the temperature on heliocentric distance is weaker for agglomerates than for blackbody, irrespective of the separation between the grains. The authors concluded that the agglomerates attain the blackbody temperature at small heliocentric distances, but this does not seem to be the case (cf. Kimura et al. 1997, 2002b). Their misleading conclusion could be attributed to simply their crude method for estimates of absorption cross sections at short wavelengths.

Mukai and Okada (2007) estimated equilibrium temperatures of BPCA ($D \approx 3$) particles consisting of silicate grains with $N = 2048$ and 16,384. They used the DDSCAT code (ver. 6.1) with the LDR method for computing the absorption cross sections of the agglomerates consisting of small constituent grains with $a_0 < 0.1\,\mu\text{m}$ and the geometric optics for large constituent grains with $a_0 \geq 10\,\mu\text{m}$. They have shown that the results between the DDA and the geometric optics could be smoothly interpolated in the range of $0.1 \leq a_0 < 10\,\mu\text{m}$, if they apply the Maxwell Garnett mixing rule to computations of absorption cross sections. Therefore, they applied the Maxwell Garnett mixing rule to calculate the temperatures of the agglomerates, irrespective of the radius of constituent grains. Levasseur-Regourd et al. (2007) took the same procedures to compute equilibrium temperatures of BPCA ($D \approx 3$) particles consisting of carbon grains with $N = 2048$ and 16,384. Because they consider agglomerates up to $a_\text{v} = 10^4\,\mu\text{m}$, the radius of constituent grains reached $a_0 \approx 40-80\,\mu\text{m}$. Because their agglomerates do not have constituent grains of $a_0 \sim 0.1\,\mu\text{m}$ except for $a_\text{v} \sim 1 - 3\,\mu\text{m}$, it does not seem plausible that their results on the size dependence of temperatures are applicable to primitive dust particles in a planetary system.

Lasue et al. (2007) used the DDSCAT code (ver. 5a10) presumably with the LDR method for computing the temperatures of BPCA ($D \approx 3$) and BCCA ($D \approx 2$) particles as a function of heliocentric distance to simulate the radial dependence of temperature inferred from zodiacal light observations. They considered agglomerates of 64 spheroidal constituent grains with $a_0 = 0.19\,\mu\text{m}$ composed of silicate or organic material, as well as compact spheres and spheroids of silicates, organic material, or amorphous carbon. Their results indicate that absorbing materials such as organics and amorphous carbon play an important role in the radial gradient of the temperature in the zodiacal light. They concluded that the radial dependence of the temperature does not strongly depend on the shape and structure of the particle, but the size and material of the particle. Their conclusion on the size dependence of equilibrium temperature was, however, not confirmed for agglomerates, as they studied the size dependence of equilibrium temperature only with compact spheres and spheroids.

7 Concluding Remarks

In situ measurements in the Solar System have revealed that the total cross sections of primitive dust particles are dominated by large agglomerates of submicron constituent grains. The mass distribution of interplanetary dust measured in situ at 1 au from the Sun shows that the cross-sectional distribution has a peak at $m \approx 3 \times 10^{-10}$ kg, which is equivalent to $a_v \approx 30\,\mu$m and $N \approx 3 \times 10^7$ (Grün et al. (1985). The predominance of large agglomerates with $m \geq 10^{-13}$ kg ($a_v \geq 2\,\mu$m) is supported by the mass distribution of dust particles in the inner coma of comet 1P/Halley measured in situ by the Dust Impact Detection System (DIDSY) onboard Giotto (McDonnell et al. 1987; Kolokolova et al. 2007). This is in harmony with the characteristic radius of $a_c \approx 20\,\mu$m ($N \approx 2 \times 10^6$) for agglomerates of submicron constituent grains with $D = 2.5$ derived from $\beta \approx 0.4$ for dust particles in the ejecta plume of comet 9P/Tempel 1 (Kobayashi et al. 2013). Because meteoritic impact is the major process of releasing dust particles from asteroids, we expect that typical dust particles ejected from asteroids are also characterized by agglomerates of $N \geq 10^6$ constituent grains. More recently, the Grain Impact Analyser and Dust Accumulator (GIADA) and the Optical, Spectroscopic, and Infrared Remote Imaging System (OSIRIS) onboard Rosetta have revealed that dust particles in the 10^{-8} to 10^{-7} kg mass bin (i.e., $a_v \sim 100-200\,\mu$m, $N \sim 10^9-10^{10}$) dominate the coma of comet 67P/Churyumov–Gerasimenko (Rotundi et al. 2015). The COmetary Secondary Ion Mass Analyser (COSIMA) onboard Rosetta has collected fluffy agglomerates with $a_c \geq 25\,\mu$m and $p > 0.5$ from the comet (Schulz et al. 2015). The predominance of large agglomerates in the total cross sections is also true of debris disks around early-type stars such as β Pic and Fomalhaut, because small particles of $a_v < 10\,\mu$m are quickly blew out by stellar radiation pressure (e.g., Kimura and Mann 1999b; Acke et al. 2012). In conclusion, it is of great importance to estimate light-scattering properties of large agglomerates consisting of million or more grains of submicrometer size.

It is worth emphasizing that available numerical techniques allow us to compute thermal emission from large ($N \geq 10^6$) agglomerates of submicron constituent grains. However, currently available numerical techniques do not allow us to properly simulate light scattering by such a large agglomerate consisting of submicron grains. As a result, radiative transfer computations and SED fittings often rely on the Henyey–Greenstein phase function without proof of its applicability to light scattering by dust agglomerates. Nevertheless, computing powers are, without doubt, increasing as the time goes by and numerical codes for calculation of light scattering by agglomerates are making steady progress. In the end, we expect that our understanding of primitive dust particles in planetary systems will make great advances once light scattering by large agglomerates of more than 10^6 submicron constituent grains becomes available.

As we have noted in Sect. 1, the current review has focused on light-scattering and thermal-emission properties of primitive dust particles that were released from planetesimals in orbit around main-sequence stars. Accordingly, we have not dealt with any subject of dust particles in protoplanetary disks, because the majority of the particles might be prestellar interstellar grains and thus not necessarily be incorporated into planetesimals. We are, however, aware of the most recent findings that the majority of protoplanetary disks around weak-line T Tauri stars (WTTS) could be regarded as young debris disks (Hardy et al. 2015). If this identification is confirmed by future studies, then we have to admit that the current review is far from complete and any coming review on this subject cannot avoid additionally discussing light-scattering and thermal-emission properties of dust particles in young debris disks around WTTSs. At the moment, we are confident that we have presented a new state-of-the-art review on light-scattering and thermal-emission properties of primitive dust particles in planetary systems.

8 Summary

We have reviewed numerical approaches to deducing light-scattering and thermal-emission properties of primitive dust particles in planetary systems from astronomical observations. Currently available information on the particles indicates that they are often agglomerates of small constituent grains whose sizes are comparable to visible wavelength. If the particles are pristine, they are composed mainly of magnesium-rich silicates, ferrous metals and sulfides, and organic refractory materials. If dust particles were subjected to metamorphism, they could be composed of hydrous minerals, metals, and to a lesser extent organic refractory materials. These unique characteristics of primitive dust particles are associated with their formation and evolution around a star whose chemical composition is essentially solar. We have demonstrated that the development of light-scattering techniques has been offering powerful tools to make a thorough investigation of dust particles in various astrophysical environments. Among numerical techniques to solve light-scattering problems, the DDA, the TMM, the GMM, and EMAs are the most common ones for practical use in astronomy. The Monte Carlo method for radiative transfer could be applied to light scattering by an ensemble of dust particles in the environment where the optical depth is on the order of unity or larger. We have shown that numerical simulations of light scattering by and thermal emission from dust agglomerates provide new state-of-the-art knowledge of primitive dust particles in planetary systems, if combined with comprehensive collections of relevant results from not only astronomical observations, but also in situ data analyses, laboratory sample analyses, laboratory analog experiments, and theoretical studies on the formation and evolution of the particles.

Acknowledgments We would like to thank anonymous reviewers for constructive comments and helpful suggestions that greatly improved the manuscript. We are grateful to Yu-lin Xu, Yasuhiko Okada, Ryo Tazaki, Sebastian Wolf, and Bruce T. Draine for useful information. We also express our gratitude to Alexander A. Kokhanovsky who encouraged us to write up a comprehensive review of this subject. H. Kimura is thankful to Grants-in-Aid for Scientific Research (#26400230, #15K05273, #23103004) from the Japan Society of the Promotion of Science.

References

Acke B, Min M, Dominik C, Vandenbussche B, Sibthorpe B, Waelkens C, Olofsson G, Degroote P, Smolders K, Pantin E, Barlow MJ, Blommaert JADL, Brandeker A, De Meester W, Dent WRF, Exter K, Di Francesco J, Fridlund M, Gear WK, Glauser AM, Greaves JS, Harvey PM, Henning T, Hogerheijde MR, Holland WS, Huygen R, Ivison RJ, Jean C, Liseau R, Naylor DA, Pilbratt GL, Polehampton ET, Regibo S, Royer P, Sicilia-Aguilar A, Swinyard BM (2012) Herschel images of Fomalhaut. An extrasolar Kuiper belt at the height of its dynamical activity. Astron Astrophys 540:A125. doi:10.1051/0004-6361/201118581

A'Hearn MF, Belton MJS, Delamere WA, Kissel J, Klaasen KP, McFadden LA, Meech KJ, Melosh HJ, Schultz PH, Sunshine JM, Thomas PC, Veverka J, Yeomans DK, Baca MW, Busko I, Crockett CJ, Collins SM, Desnoyer M, Eberhardy CA, Ernst CM, Farnham TL, Feaga L, Groussin O, Hampton D, Ipatov SI, Li J-Y, Lindler D, Lisse CM, Mastrodemos N, Owen Jr WM, Richardson JE, Wellnitz DD, White RL (2005) Deep Impact: excavating comet Tempel 1. Science 310:258–264. doi:10.1126/science.1118923

Altobelli N, Postberg F, Fiege K, Trieloff M, Kimura H, Sterken VJ, Hsu H-W, Hillier J, Khawaya N, Moragas-Klostermeyer G, Blum J, Burton M, Srama R, Kempf S, Grün E (2016) The Cosmic Dust Analyzer onboard Cassini measures flux and composition of interstellar dust at Saturn. Science, in press

Arakawa M, Saiki T, Wada K, Kadono T, Takagi Y, Shirai K, Okamoto C, Yano H, Hayakawa M, Nakazawa S, Hirata N, Kobayashi M, Michel P, Jutzi M, Imamura H, Ogawa K, Iijima Y, Honda R, Ishibashi K, Hayakawa H, Sawada H (2013) Small Carry-on Impactor (SCI): its scientific purpose, operation, and observation plan in Hayabysa-2 mission. In: 44th Lunar and Planetary Science Conference, p 1904

Artymowicz P, Clampin M (1997) Dust around main-sequence stars: nature or nurture by the interstellar medium? Astrophys J 490:863–878. doi:10.1086/304889

Augereau JC, Lagrange AM, Mouillet D, Papaloizou JCB, Grorod PA (1999) On the HR 4796 A circumstellar disk. Astron Astrophys 348:557–569

Backman DE, Gillett FC, Witteborn FC (1992) Infrared observations and thermal models of the β Pictoris disk. Astrophys J 385:670–679. doi:10.1086/170973

Bemporad A, Poletto G, Raymond JC (2006) Evidence for pyroxene dust grains in C/2001 C2 sungrazing comet. Adv Space Res 38(9):1972–1975. doi:10.1016/j.asr.2005.10.005

Bemporad A, Poletto G, Raymond J, Giordano S (2007) A review of SOHO/UVCS observations of sungrazing comets. Planet Space Sci 55:1021–1030. doi:10.1016/j.pss.2006.11.013

Bertini I, Thomas N, Barbieri C (2007) Modeling of the light scattering properties of cometary dust using fractal aggregates. Astron Astrophys 461:351–364. doi:10.1051/0004-6361:20065461

Biesecker DA, Lamy P, St Cyr OC, Llebaria A, Howard RA (2002) Sungrazing comets discovered with the SOHO/LASCO coronagraphs 1996–1998. Icarus 157:323–348. doi:10.1006/icar.2002.6827

Bohren CF, Huffman DR (1983) Absorption and scattering of light by small particles. Wiley-Interscience, New York

Bregman JD, Campins H, Witteborn FC, Wooden DH, Rank DM, Allamandola LJ, Cohen M, Tielens AGGM (1987) Airborne and groundbased spectrophotometry of comet P/Halley from 5–13 micrometers. Astron Astrophys 187:616–620

Brownlee DE (1985) Cosmic dust: collection and research. Annu Rev Earth Planet Sci 13:147–173. doi:10.1146/annurev.ea.13.050185.001051

Brownlee DE, Tomandl DA, Hodge PW (1976) Extraterrestrial particles in the stratosphere. In: Elsässer H, Fechtig H (eds) Interplanetary dust and zodiacal light. Springer, Berlin, pp 279–283

Brunetto R, Borg J, Dartois E, Rietmeijer FJM, Grossemy F, Sandt C, Le Sergeant d'Hendecourt L, Rotundi A, Dumas P, Djouadi Z, Jamme F (2011) Mid-IR, Far-IR, Raman micro-spectroscopy, and FESEM-EDX study of IDP L2021C5: clues to its origin. Icarus 212:896–910. doi:10.1016/j.icarus.2011.01.038

Campins H, Ryan EV (1989) The identification of crystalline olivine in cometary silicates. Astrophys J 341:1059–1066. doi:10.1086/167563

Chen CH, Li A, Bohac C, Kim KH, Watson DM, Van Cleve J, Houck J, Stapelfeldt K, Werner MW, Rieke G, Su K, Marengo M, Backman D, Beichman C, Fazio G (2007) The dust and gas around β Pictoris. Astrophys J 666:466–474. doi:10.1086/519989

Collinge MJ, Draine BT (2004) Discrete-dipole approximation with polarizabilities that account for both finite wavelength and target geometry. J Opt Soc Am A 21(10):2023–2028. doi:10.1364/JOSAA.21.002023

Day KL (1975) Measured extinction of small olivine spheres. Astrophys J 199:660–662. doi:10.1086/153734

Debes JH, Weinberger AJ, Schneider G (2008) Complex organic materials in the circumstellar disk of HR 4796A. Astrophys J 673:L191–L194. doi:10.1086/527546

Dominik C, Tielens AGGM (1997) The physics of dust coagulation and the structure of dust aggregates in space. Astrophys J 480:647–673. doi:10.1086/303996

Donaldson JK, Lebreton J, Roberge A, Augereau J-C, Krivov AV (2013) Modeling the HD 32297 debris disk with far-infrared Herschel data. Astrophys J 772:17 (10pp). doi:10.1088/0004-637X/772/1/17

Dorschner J, Friedemann C, Gürtler J (1978) Laboratory spectra of phyllosilicates and the interstellar 10-micrometer absorption band. Astron Nachr 299:269–282

Draine BT (1988) The discrete-dipole approximation and its application to interstellar graphite grains. Astrophys J 333:848–872. doi:10.1086/166795

Draine BT, Flatau PJ (1994) Discrete-dipole approximation for scattering calculations. J Opt Soc Am A 11(4):1491–1499. doi:10.1364/JOSAA.11.001491

Draine BT, Goodman J (1993) Beyond Clausius-Mossotti: wave propagation on a polarizable point lattice and the discrete dipole approximation. Astrophys J 405:685–697. doi:10.1086/172396

Fitzgerald MP, Kalas PG, Duchêne G, Pinte C, Graham JR (2007) The AU Microscopii debris disk: multiwavelength imaging and modeling. Astrophys J 670:536–556. doi:10.1086/521344

Fomenkova MN (1999) On the organic refractory component of cometary dust. Space Sci Rev 90:109–114. doi:10.1023/A:1005237828783

Gehrz RD, Ney EP (1992) 0.7- to 23-μm photometric observations of P/Halley 1986 III and six recent bright comets. Icarus 100:162–186. doi:10.1016/0019-1035(92)90027-5

Golimowski DA, Ardila DR, Krist JE, Clampin M, Ford HC, Illingworth GD, Bartko F, Benítez N, Blakeslee JP, Bouwens RJ, Bradley LD, Broadhurst TJ, Brown RA, Burrows CJ, Cheng ES, Cross NJG, Demarco R, Feldman PD, Franx M, Goto T, Gronwall C, Hartig GF, Holden BP, Homeier NL, Infante L, Jee MJ, Kimble RA, Lesser MP, Martel AR, Mei S, Menanteau F, Meurer GR, Miley GK, Motta V, Postman M, Rosati P, Sirianni M, Sparks WB, Tran HD, Tsvetanov ZI, White RL, Zheng W, Zirm AW (2006) Hubble Space Telescope ACS multiband coronagraphic imaging of the debris disk around β Pictoris. Astron J 131:3109–3130. doi:10.1086/503801

Golimowski DA, Krist JE, Stapelfeldt KR, Chen CH, Ardila DR, Bryden G, Clampin M, Ford HC, Illingworth GD, Plavchan P, Rieke GH, Su KYL (2011) Hubble and Spitzer Space Telescope

observations of the debris disk around the nearby K dwarf HD 92945. Astron J 142:30 (16pp). doi:10.1088/0004-6256/142/1/30
Graham JR, Kalas PG, Matthews BC (2007) The signature of primordial grain growth in the polarized light of the AU Microscopii debris disk. Astrophys J 654:595–605. doi:10.1086/509318
Greenberg JM (1984) Evolution of interstellar grains. In: Wolstencroft RD, Greenberg JM (eds) Proceedings of the workshop on laboratory and observational infrared spectra of interstellar dust. Royal Observatory, Edinburgh, pp 1–25
Greenberg JM, Hage JI (1990) From interstellar dust to comets: a unification of observational constraints. Astrophys J 361:260–274. doi:10.1086/169191
Greenberg JM, Li A (1998) From interstellar dust to comets: the extended CO source in comet Halley. Astron Astrophys 332:374–384
Grigorieva A, Artymowicz P, Thébault P (2007) Collisional dust avalanches in debris discs. Astron Astrophys 461:537–549. doi:10.1051/0004-6361:20065210
Grossman L, Steele IM (1976) Amoeboid olivine aggregates in the Allende meteorite. Geochim Cosmochim Acta 40:149–155. doi:10.1016/0016-7037(76)90172-1
Grün E, Zook HA, Fechtig H, Giese RH (1985) Collisional balance of the meteoritic complex. Icarus 62:244–272. doi:10.1016/0019-1035(85)90121-6
Gustafson BÅS, Kolokolova L (1999) A systematic study of light scattering by aggregate particles using the microwave analog technique: angular and wavelength dependence of intensity and polarization. J Geophys Res 104:31711–31720. doi:10.1029/1999JD900327
Gutkowicz-Krusin D, Draine BT (2004) Propagation of electromagnetic waves on a rectangular lattice of polarizable points. http://arxiv.org/abs/astro-ph/0403082
Hage JI, Greenberg JM (1990) A model for the optical properties of porous grains. Astrophys J 361:251–259. doi:10.1086/169190
Hanner MS (1980) On the albedo of the interplanetary dust. Icarus 43:373–380. doi:10.1016/0019-1035(80)90181-5
Hanner MS, Giese RH, Weiss K, Zerull R (1981) On the definition of albedo and application to irregular particles. Astron Astrophys 104:42–46
Hardy A, Caceres C, Schreiber MR, Cieza L, Alexander RD, Canovas H, Williams JP, Wahhaj Z, Menard F (2015) Probing the final stages of protoplanetary disk evolution with ALMA. Astron Astrophys 583:A66. doi:10.1051/0004-6361/201526504
Harker DE, Woodward CE, Wooden DH, Fisher RS, Trujillo CA (2007) Gemini-N mid-IR observations of the dust properties of the ejecta excavated from comet 9P/Tempel 1 during Deep Impact. Icarus 190:432–453. doi:10.1016/j.icarus.2007.03.008
Henyey LG, Greenstein JL (1941) Diffuse radiation in the galaxy. Astrophys J 93:70–83. doi:10.1086/144246
Ishiguro M, Yang H, Usui F, Pyo J, Ueno M, Ootsubo T, Kwon SM, Mukai T (2013) High-resolution imaging of the Gegenschein and the geometric albedo of interplanetary dust. Astrophys J 767:75 (13pp). doi:10.1088/0004-637X/767/1/75
Ishimoto H, Mann I (1999) Modeling the particle mass distribution within 1 AU of the Sun. Planet Space Sci 47:225–232. doi:10.1016/S0032-0633(98)00083-X
Jenniskens P (1993) Optical constants of organic refractory residue. Astron Astrophys 274:653–661
Jenniskens P, Baratta GA, Kouchi A, de Groot MS, Greenberg JM, Strazzulla G (1993) Carbon dust formation on interstellar grains. Astron Astrophys 273:583–600
Jessberger EK (1999) Rocky cometary particulates: their elemental, isotopic and mineralogical ingredients. Space Sci Rev 90:91–97. doi:10.1023/A:1005233727874
Jessberger EK, Christoforidis A, Kissel J (1988) Aspects of the major element composition of Halley's dust. Nature 332:691–695. doi:10.1038/332691a0
Keller LP, Messenger S (2011) On the origins of GEMS grains. Geochim Cosmochim Acta 75:5336–5365. doi:10.1016/j.gca.2011.06.040

Kimura H (2001) Light-scattering properties of fractal aggregates: numerical calculations by a superposition technique and the discrete-dipole approximation. J Quant Spectrosc Radiat Transfer 70:581–594. doi:10.1016/S0022-4073(01)00031-0

Kimura H (2013) The reincarnation of interstellar dust: the importance of organic refractory material in infrared spectra of cometary comae and circumstellar disks. Astrophys J 775: L18 (5pp). doi:10.1088/2041-8205/775/1/L18

Kimura H (2014) The organic-rich carbonaceous component of dust aggregates in circumstellar disks: effects of its carbonization on infrared spectral features of its magnesium-rich olivine counterpart. Icarus 232:133–140. doi:10.1016/j.icarus.2014.01.009

Kimura H, Mann I (1998) Radiation pressure cross section for fluffy aggregates. J Quant Spectrosc Radiat Transfer 60:425–438. doi:10.1016/S0022-4073(98)00017-X

Kimura H, Mann I (1999a) Radiation pressure on porous micrometeoroids. In: Baggaley WJ, Porubčan V (eds) Meteoroids 1998. Astronomical Institute of the Slovak Academy of Sciences, Bratislava, pp 283–286

Kimura H, Mann I (1999b) Radiation pressure on dust aggregates in circumstellar disks. Phys Chem Earth C Solar Terr Planet Sci 24(5):561–566. doi:10.1016/S1464-1917(99)00092-6

Kimura H, Mann I (2004) Light scattering by large clusters of dipoles as an analog for cometary dust aggregates. J Quant Spectrosc Radiat Transfer 89:155–164. doi:10.1016/j.jqsrt.2004.05.019

Kimura H, Ishimoto H, Mukai T (1997) A study on solar dust ring formation based on fractal dust models. Astron Astrophys 326:263–270

Kimura H, Okamoto H, Mukai T (2002a) Radiation pressure and the Poynting-Robertson effect for fluffy dust particles. Icarus 157:349–361. doi:10.1006/icar.2002.6849

Kimura H, Mann I, Biesecker DA, Jessberger EK (2002b) Dust grains in the comae and tails of sungrazing comets: modeling of their mineralogical and morphological properties. Icarus 159:529–541. doi:10.1006/icar.2002.6940

Kimura H, Mann I, Jessberger EK (2003a) Elemental abundances and mass densities of dust and gas in the Local Interstellar Cloud. Astrophys J 582:846–858. doi:10.1086/344691

Kimura H, Mann I, Jessberger EK (2003b) Composition, structure, and size distribution of dust in the Local Interstellar Cloud. Astrophys J 583:314–321. doi:10.1086/345102

Kimura H, Kolokolova L, Mann I (2003c) Optical properties of cometary dust. Astron Astrophys 407:L5–L8. doi:10.1051/0004-6361:20030967

Kimura H, Kolokolova L, Mann I (2006) Light scattering by cometary dust numerically simulated with aggregate particles consisting of identical spheres. Astron Astrophys 449:1243–1254. doi:10.1051/0004-6361:20041783

Kimura H, Chigai T, Yamamoto T (2008) Mid-infrared spectra of cometary dust: the evasion of its silicate mineralogy. Astron Astrophys 482:305–307. doi:10.1051/0004-6361:20078778

Kimura H, Chigai T, Yamamoto T (2009) Infrared spectra of dust aggregates in cometary comae: calculation with olivine formed by exothermic chemical reactions. Astrophys J 690:1590–1596. doi:10.1088/0004-637X/690/2/1590

Kimura H, Senshu H, Wada K (2014) Electrostatic lofting of dust aggregates near the terminator of airless bodies and its implication for the formation of exozodiacal disks. Planet Space Sci 100:64–72. doi:10.1016/j.pss.2014.03.017

Kirchschlager F, Wolf S (2013) Porous dust grains in debris disks. Astron Astrophys 552:A54. doi:10.1051/0004-6361/201220486

Knacke RF, Fajardo-Acosta SB, Telesco CM, Hackwell JA, Lynch DK, Russell RW (1993) The silicates in the disk of β Pictoris. Astrophys J 418:440–450. doi:10.1086/173405

Kobayashi H, Kimura H, Yamamoto S (2013) Dust mantle of comet 9P/Tempel 1: dynamical constraints on physical properties. Astron Astrophys 550:A72. doi:10.1051/0004-6361/201220464

Köhler M, Minato T, Kimura H, Mann I (2007) Radiation pressure force acting on cometary aggregates. Adv Space Res 40(2):266–271. doi:10.1016/j.asr.2007.05.044

Köhler M, Mann I, Li A (2008) Complex organic materials in the HR 4796A disk? Astrophys J 686:L95–L98. doi:10.1086/592961

Kolokolova L, Kimura H (2010a) Comet dust as a mixture of aggregates and solid particles: model consistent with ground-based and space-mission results. Earth Planets Space 62:17–21. doi:10.5047/eps.2008.12.001

Kolokolova L, Kimura H (2010b) Effects of electromagnetic interaction in the polarization of light scattered by cometary and other types of cosmic dust. Astron Astrophys 513:A40. doi:10.1051/0004-6361/200913681

Kolokolova L, Mackowski D (2012) Polarization of light scattered by large aggregates. J Quant Spectrosc Radiat Transfer 113:2567–2572. doi:10.1016/j.jqsrt.2012.02.002

Kolokolova L, Kimura H, Mann I (2005) Characterization of dust particles using photopolarimetric data: example of cometary dust. In: Videen G, Yatskiv Y, Mishchenko M (eds) Photopolarimetry in remote sensing. Kluwer Academic Publishers, Dordrecht, pp 431–454. doi:10.1007/1-4020-2368-5_20

Kolokolova L, Kimura H, Kiselev N, Rosenbush V (2007) Two different evolutionary types of comets proved by polarimetric and infrared properties of their dust. Astron Astrophys 463:1189–1196. doi:10.1051/0004-6361:20065069

Kozasa T, Blum J, Mukai T (1992) Optical properties of dust aggregates. I. wavelength dependence. Astron Astrophys 263:423–432

Kozasa T, Blum J, Okamoto H, Mukai T (1993) Optical properties of dust aggregates. II. angular dependence of scattered light. Astron Astrophys 276:278–288

Krist JE, Ardila DR, Golimowski DA, Clampin M, Ford HC, Illingworth GD, Hartig GF, Bartko F, Benítez N, Blakeslee JP, Bouwens RJ, Bradley LD, Broadhurst TJ, Brown RA, Burrows CJ, Cheng ES, Cross NJG, Demarco R, Feldman PD, Franx M, Goto T, Gronwall C, Holden B, Homeier N, Infante L, Kimble RA, Lesser MP, Martel AR, Mei S, Menanteau F, Meurer GR, Miley GK, Motta V, Postman M, Rosati P, Sirianni M, Sparks WB, Tran HD, Tsvetanov ZI, White RL, Zheng W (2005) Hubble Space Telescope Advanced Camera for surveys coronagraphic imaging of the AU Microscopii debris disk. Astron J 129:1008–1017. doi:10.1086/426755

Krist JE, Stapelfeldt KR, Bryden G, Rieke GH, Su KYL, Chen CC, Beichman CA, Hines DC, Rebull LM, Tanner A, Trilling DE, Clampin M, Gáspár A (2010) HST and Spitzer observations of the HD 207129 debris ring. Astron J 140:1051–1061. doi:10.1088/0004-6256/140/4/1051

Krivov AV, Mann I, Krivova NA (2000) Size distributions of dust in circumstellar debris discs. Astron Astrophys 362:1127–1137

Lamy PL, Malburet P, Llebaria A, Koutchmy S (1989) Comet P/Halley at a heliocentric preperihelion distance of 2.6 AU: jet activity and properties of the dust coma. Astron Astrophys 222:316–322

Lasue J, Levasseur-Regourd AC (2006) Porous irregular aggregates of sub-micron sized grains to reproduce cometary dust light scattering observations. J Quant Spectrosc Radiat Transfer 100:220–236. doi:10.1016/j.jqsrt.2005.11.040

Lasue J, Levasseur-Regourd AC, Fray N, Cottin H (2007) Inferring the interplanetary dust properties from remote observations and simulations. Astron Astrophys 473:641–649. doi:10.1051/0004-6361:20077623

Lasue J, Levasseur-Regourd AC, Hadamcik E, Alcouffe G (2009) Cometary dust properties retrieved from polarization observations: application to C/1995 O1 Hale-Bopp and 1P/Halley. Icarus 199:129–144. doi:10.1016/j.icarus.2008.09.008

Lebreton J, Augereau J-C, Thi W-F, Roberge A, Donaldson J, Schneider G, Maddison ST, Ménard F, Riviere-Marichalar P, Mathews GS, Kamp I, Pinte C, Dent WRF, Barrado D, Duchêne G, Gonzalez J-F, Grady CA, Meeus G, Pantin E, Williams JP, Woitke P (2012) An icy Kuiper belt around the young solar-type star HD 181327. Astron Astrophys 539:A17. doi:10.1051/0004-6361/201117714

Lebreton J, Beichman C, Augereau JC (2014) Detailed models of a sample of debris disks: from Herschel, KIN and Spitzer to the JWST. In: Ballet J, Bournaud F, Martins F, Monier R, Reylé C (eds) Proceedings of the annual meeting of the French Society of Astronomy and Astrophysics (SF2A 2014), pp 193–196

Levasseur-Regourd AC, Hadamcik E, Renard JB (1996) Evidence for two classes of comets from their polarimetric properties at large phase angles. Astron Astrophys 313:327–333

Levasseur-Regourd AC, Cabane M, Worms JC, Haudebourg V (1997) Physical properties of dust in the solar system: relevance of a computational approach and of measurements under microgravity conditions. Adv Space Res 20(8):1585–1594. doi:10.1016/S0273-1177(97)00816-8

Levasseur-Regourd AC, Mukai T, Lasue J, Okada Y (2007) Physical properties of cometary and interplanetary dust. Planet Space Sci 55:1010–1020. doi:10.1016/j.pss.2006.11.014

Li A, Greenberg JM (1998) A comet dust model for the β Pictoris disk. Astron Astrophys 331:291–313

Li A, Lunine JI (2003) Modeling the infrared emission from the HR 4796A disk. Astrophys J 590:368–378. doi:10.1086/374865

Li A, Lunine JI, Bendo GJ (2003) Modeling the infrared emission from the ε Eridani disk. Astrophys J 598:L51–L54. doi:10.1086/380495

Lumme K, Penttilä A (2011) Model of light scattering by dust particles in the solar system: applications to cometary comae and planetary regoliths. J Quant Spectrosc Radiat Transfer 112:1658–1670. doi:10.1016/j.jqsrt.2011.01.016

Lumme K, Rahola J, Hovenier JW (1997) Light scattering by dense clusters of spheres. Icarus 126:455–469. doi:10.1006/icar.1996.5650

Mackowski DW (2002) Discrete dipole moment method for calculation of the T matrix for nonspherical particles. J Opt Soc Am A 19(5):881–893. doi:10.1364/JOSAA.19.000881

Mackowski DW, Mishchenko MI (1996) Calculation of the T matrix and the scattering matrix for ensembles of spheres. J Opt Soc Am A 13(11):2266–2278. doi:10.1364/JOSAA.13.002266

Mackowski DW, Mishchenko MI (2011) A multiple sphere T-matrix Fortran code for use on parallel computer clusters. J Quant Spectrosc Radiat Transfer 112:2182–2192. doi:10.1016/j.jqsrt.2011.02.019

Mann I, Okamoto H, Mukai T, Kimura H, Kitada Y (1994) Fractal aggregate analogues for near solar dust properties. Astron Astrophys 291:1011–1018

Mann I, Kimura H, Kolokolova L (2004) A comprehensive model to describe light scattering properties of cometary dust. J Quant Spectrosc Radiat Transfer 89:291–301. doi:10.1016/j.jqsrt.2004.05.029

Mathis JS, Whiffen G (1989) Composite interstellar grains. Astrophys J 341:808–822. doi:10.1086/167538

McDonnell JAM, Alexander WM, Burton WM, Bussoletti E, Evans GC, Evans ST, Firth JG, Grard RJL, Green SF, Grun E, Hanner MS, Hughes DW, Igenbergs E, Kissel J, Kuczera H, Lindblad BA, Langevin Y, Mandeville J-C, Nappo S, Pankiewicz GSA, Perry CH, Schwehm GH, Sekanina Z, Stevenson TJ, Turner RF, Weishaupt U, Wallis MK, Zarnecki JC (1987) The dust distribution within the inner coma of comet P/Halley 1982i: encounter by Giotto's impact detectors. Astron Astrophys 187:719–741

Meakin P (1984) Effects of cluster trajectories on cluster-cluster aggregation: a comparison of linear and Brownian trajectories in two- and three-dimensional simulations. Phys Rev A 29:997–999. doi:10.1103/PhysRevA.29.997

Meakin P, Donn B (1988) Aerodynamic properties of fractal grains: implications for the primordial solar nebula. Astrophys J 329:L39–L41. doi:10.1086/185172

Minato T, Köhler M, Kimura H, Mann I, Yamamoto T (2006) Momentum transfer to fluffy dust aggregates from stellar winds. Astron Astrophys 452:701–707. doi:10.1051/0004-6361:20054774

Moreno F, Muñoz O, Guirado D, Vilaplana R (2007) Comet dust as a size distribution of irregularly shaped, compact particles. J Quant Spectrosc Radiat Transfer 106:348–359. doi:10.1016/j.jqsrt.2007.01.023

Mukai T, Okada Y (2007) Optical properties of large aggregates. In: Workshop on dust in planetary systems, SP-643. ESA Publications Division, Noordwijk, pp 157–160

Mukai T, Ishimoto H, Kozasa T, Blum J, Greenberg JM (1992) Radiation pressure forces of fluffy porous grains. Astron Astrophys 262:315–320

Nagdimunov L, Kolokolova L, Wolff M, A'Hearn MF, Farnham TL (2014) Properties of comet 9P/Tempel 1 dust immediately following excavation by Deep Impact. Planet Space Sci 100:73–78. doi:10.1016/j.pss.2014.05.018

Nakamura R (1998) Optical properties of dust aggregates in the disk of Beta Pictoris. Earth Planet Space 50:587–593. doi:10.1186/BF03352152

Nakamura R, Okamoto H (1999) Optical properties of fluffy aggregates as analogue of interplanetary dust particles. Adv Space Res 23(7):1209–1212. doi:10.1016/S0273-1177(99)00185-4

Noguchi T, Ohashi N, Tsujimoto S, Mitsunari T, Bradley JP, Nakamura T, Toh S, Stephan T, Iwata N, Imae N (2015) Cometary dust in Antarctic ice and snow: past and present chondritic porous micrometeorites preserved on the Earth's surface. Earth Planet Sci Lett 410:1–11. doi:10.1016/j.epsl.2014.11.012

Okada Y (2008) Efficient numerical orientation averaging of light scattering properties with a quasi-Monte-Carlo method. J Quant Spectrosc Radiat Transfer 109:1719–1742. doi:10.1016/j.jqsrt.2008.01.002

Okamoto H (1995) Light scattering by clusters: the a1-term method. Opt Rev 2(6):407–412

Okamoto H, Xu Y-L (1998) Light scattering by irregular interplanetary dust particles. Earth Planet Space 50:577–585. doi:10.1186/BF03352151

Okamoto H, Mukai T, Kozasa T (1994) The 10 μm feature of aggregates in comets. Planet Space Sci 42:643–649. doi:10.1016/0032-0633(94)90041-8

Ossenkopf V (1993) Dust coagulation in dense molecular clouds: the formation of fluffy aggregates. Astron Astrophys 280:617–646

Penttilä A, Lumme K (2011) Optimal cubature on the sphere and other orientation averaging schemes. J Quant Spectrosc Radiat Transfer 112:1741–1746. doi:10.1016/j.jqsrt.2011.02.001

Penttilä A, Zubko E, Lumme K, Muinonen K, Yurkin MA, Draine B, Rahola J, Hoekstra AG, Shkuratov Y (2007) Comparison between discrete dipole implementations and exact techniques. J Quant Spectrosc Radiat Transfer 106:417–436. doi:10.1016/j.jqsrt.2007.01.026

Petrova EV, Tishkovets VP (2011) Light scattering by aggregates of varying porosity and the opposition phenomena observed in the low-albedo particulate media. J Quant Spectrosc Radiat Transfer 112:2226–2233. doi:10.1016/j.jqsrt.2011.01.011

Petrova EV, Jockers K, Kiselev NN (2000) Light scattering by aggregates with sizes comparable to the wavelength: an application to cometary dust. Icarus 148:526–536. doi:10.1006/icar.2000.6504

Petrova EV, Jockers K, Kiselev NN (2001a) Light scattering by aggregate particles comparable in size to wavelength: application to cometary dust. Sol Syst Res 35:57–69. doi:10.1023/A:1005267824582

Petrova EV, Jockers K, Kiselev NN (2001b) A negative branch of polarization for comets and atmosphereless celestial bodies and the light scattering by aggregate particles. Sol Syst Res 35:390–399. doi:10.1023/A:1012304321440

Petrova EV, Tishkovets VP, Jockers K (2004) Polarization of light scattered by solar system bodies and the aggregate model of dust particles. Sol Syst Res 38:309–324. doi:10.1023/B:SOLS.0000037466.32514.fe

Pinte C, Ménard F, Duchêne G, Bastien P (2006) Monte Carlo radiative transfer in protoplanetary disks. Astron Astrophys 459:797–804. doi:10.1051/0004-6361:20053275

Purcell EM, Pennypacker CR (1973) Scattering and absorption of light by nonspherical dielectric grains. Astrophys J 186:705–714. doi:10.1086/152538

Robertson HP (1937) Dynamical effects of radiation in the solar system. Mon Not R Astron Soc 97:423–437. doi:10.1093/mnras/97.6.423

Rodigas TJ, Debes JH, Hinz PM, Mamajek EE, Pecaut MJ, Currie T, Bailey V, Defrere D, De Rosa RJ, Hill JM, Leisenring J, Schneider G, Skemer AJ, Skrutskie M, Vaitheeswaran V, Ward-Duong K (2014) Does the debris disk around HD 32297 contain cometary grains? Astrophys J 783:21 (12pp). doi:10.1088/0004-637X/783/1/21

Rotundi A, Sierks H, Della Corte V, Fulle M, Gutierrez PJ, Lara L, Barbieri C, Lamy PL, Rodrigo R, Koschny D, Rickman H, Keller HU, López Moreno JJ, Accolla M, Agarwal J,

A'Hearn MF, Altobelli N, Angrilli F, Barucci MA, Bertaux J-L, Bertini I, Bodewits D, Bussoletti E, Colangeli L, Cosi M, Cremonese G, Crifo J-F, Da Deppo V, Davidsson B, Debei S, De Cecco M, Esposito F, Ferrari M, Fornasier S, Giovane F, Gustafson B, Green SF, Groussin O, Grün E, Güttler C, Herranz ML, Hviid SF, Ip W, Ivanovski S, Jerónimo JM, Jorda L, Knollenberg J, Kramm R, Kührt E, Küppers M, Lazzarin M, Leese MR, López-Jiménez AC, Lucarelli F, Lowry SC, Marzari F, Epifani EM, McDonnell JAM, Mennella V, Michalik H, Molina A, Morales R, Moreno F, Mottola S, Naletto G, Oklay N, Ortiz JL, Palomba E, Palumbo P, Perrin J-M, Rodríguez J, Sabau L, Snodgrass C, Sordini R, Thomas N, Tubiana C, Vincent J-B, Weissman P, Wenzel K-P, Zakharov V, Zarnecki JC (2015) Dust measurements in the coma of comet 67P/Churyumov-Gerasimenko inbound to the Sun. Science 347: aaa3905. doi:10.1126/science.aaa3905

Saija R, Iatì MA, Giusto A, Borghese F, Denti P, Aiello S, Cecchi-Pestellini C (2003) Radiation pressure cross-sections of fluffy interstellar grains. Mon Not R Astron Soc 341:1239–1245. doi:10.1046/j.1365-8711.2003.06490.x

Schneider G, Smith BA, Becklin EE, Koerner DW, Meier R, Hines DC, Lowrance PJ, Terrile RJ, Thompson RI, Rieke M (1999) NICMOS imaging of the HR 4796A circumstellar disk. Astrophys J 513:L127–L130. doi:10.1086/311921

Schulz R, Hilchenbach M, Langevin Y, Kissel J, Silen J, Briois C, Engrand C, Hornung K, Baklouti D, Bardyn A, Cottin H, Fischer H, Fray N, Godard M, Lehto H, Le Roy L, Merouane S, Orthous-Daunay F-R, Paquette J, Rynö J, Siljeström S, Stenzel O, Thirkell L, Varmuza K, Zaprudin B (2015) Comet 67P/Churyumov-Gerasimenko sheds dust coat accumulated over the past four years. Nature 518:216–218. doi:10.1038/nature14159

Sekanina Z (2000) Solar and Heliospheric Observatory sungrazing comets with prominent tails: evidence on dust-production peculiarities. Astrophys J 545:L69–L72. doi:10.1086/317336

Sekanina Z, Chodas PW (2012) Comet C/2011 W3 (Lovejoy): orbit determination, outbursts, disintegration of nucleus, dust-tail morphology, and relationship to new cluster of bright sungrazers. Astrophys J 757:127 (33pp). doi:10.1088/0004-637X/757/2/127

Shalima P, Wada K, Kimura H (2015) Ejecta curtain radiative transfer modeling for probing its geometry and dust optical properties. Planet Space Sci 116:39–47. doi:10.1016/j.pss.2015.03.017

Shen Y, Draine BT, Johnson ET (2009) Modeling porous dust grains with ballistic aggregates. II. light scattering properties. Astrophys J 696:2126–2137. doi:10.1088/0004-637X/696/2/2126

Silsbee K, Draine BT (2016) Radiation pressure on fluffy submicron-sized grains. Astrophys J 818:133 (9pp). doi:10.3847/0004-637X/818/2/133

Singham MK, Singham SB, Salzman GC (1986) The scattering matrix for randomly oriented particles. J Chem Phys 85(7):3807–3815. doi:10.1063/1.450901

Tazaki R, Nomura H (2015) Outward motion of porous dust aggregates by stellar radiation pressure in protoplanetary disks. Astrophys J 799:119 (9pp). doi:10.1088/0004-637X/799/2/119

Thompson WT (2015) Linear polarization measurements of Comet C/2011 W3 (Lovejoy) from STEREO. Icarus 261:122–132. doi:10.1016/j.icarus.2015.08.018

Tiscareno MS, Mitchell CJ, Murray CD, Di Nino D, Hedman MM, Schmidt J, Burns JA, Cuzzi JN, Porco CC, Beurle K, Evans MW (2013) Observations of ejecta clouds produced by impacts onto Saturn's rings. Science 340:460–464. doi:10.1126/science.1233524

Van de Hulst HC (1957) Light scattering by small particles. Dover, New York

Videen G, Muinonen K (2015) Light-scattering evolution from particles to regolith. J Quant Spectrosc Radiat Transfer 150:87–94. doi:10.1016/j.jqsrt.2014.05.019

Voshchinnikov NV, Il'in VB, Henning T (2005) Modelling the optical properties of composite and porous interstellar grains. Astron Astrophys 429:371–381. doi:10.1051/0004-6361:200400081

Wada K, Tanaka H, Suyama T, Kimura H, Yamamoto T (2008) Numerical simulation of dust aggregate collisions. II. compression and disruption of three-dimensional aggregates in head-on collisions. Astrophys J 677:1296–1308. doi:10.1086/529511

Weidenschilling SJ (1997) The origin of comets in the solar nebula: a unified model. Icarus 127:290–306. doi:10.1006/icar.1997.5712

Weidenschilling SJ, Donn B, Meakin P (1989) The physics of planetesimal formation. In: Weaver HA, Danly L, Fall S (eds) The formation and evolution of planetary systems. Cambridge University Press, Cambridge, pp 131–150

Weinberg JL, Beeson DE (1976) Polarization reversal in the tail of comet Ikeya-Seki (1965 VIII). Astron Astrophys 48:151–153

West RA (1991) Optical properties of aggregate particles whose outer diameter is comparable to the wavelength. Appl Opt 30(36):5316–5324. doi:10.1364/AO.30.005316

Westphal AJ, Stroud RM, Bechtel HA, Brenker FE, Butterworth AL, Flynn GJ, Frank DR, Gainsforth Z, Hillier JK, Postberg F, Simionovici AS, Sterken VJ, Nittler LR, Allen C, Anderson D, Ansari A, Bajt S, Bastien RK, Bassim N, Bridges J, Brownlee DE, Burchell M, Burghammer M, Changela H, Cloetens P, Davis AM, Doll R, Floss C, Grün E, Heck PR, Hoppe P, Hudson B, Huth J, Kearsley A, King AJ, Lai B, Leitner J, Lemelle L, Leonard A, Leroux H, Lettieri R, Marchant W, Ogliore R, Ong WJ, Price MC, Sandford SA, Sans Tresseras J-A, Schmitz S, Schoonjans T, Schreiber K, Silversmit G, Solé VA, Srama R, Stadermann F, Stephan T, Stodolna J, Sutton S, Trieloff M, Tsou P, Tyliszczak T, Vekemans B, Vincze L, Von Korff J, Wordsworth N, Zevin D, Zolensky ME, 30714 Stardust@home dusters (2014) Evidence for interstellar origin of seven dust particles collected by the Stardust spacecraft. Science 345:786–791. doi:10.1126/science.1252496

Wilck M, Mann I (1996) Radiation pressure forces on "typical" interplanetary dust grains. Planet Space Sci 44:493–499. doi:10.1016/0032-0633(95)00151-4

Wolff MJ, Clayton GC, Gibson SJ (1998) Modeling composite and fluffy grains. II. porosity and phase functions. Astrophys J 503:815–830. doi:10.1086/306029

Wooden DH, Harker DE, Woodward CE, Butner HM, Koike C, Witteborn FC, McMurtry CW (1999) Silicate mineralogy of the dust in the inner coma of comet C/1995 01 (Hale-Bopp) pre- and post-perihelion. Astrophys J 517:1034–1058. doi:10.1086/307206

Wooden DH, Woodward CE, Harker DE (2004) Discovery of crystalline silicates in comet C/2001 Q4 (NEAT). Astrophys J 612:L77–L80. doi:10.1086/424593

Wurm G, Blum J (1998) Experiments on preplanetary dust aggregation. Icarus 132:125–136. doi:10.1006/icar.1998.5891

Wyatt Jr SP, Whipple FL (1950) The Poynting-Robertson effect on meteor orbits. Astrophys J 111:134–141. doi:10.1086/145244

Xing Z, Hanner MS (1996) Modelling the temperature of cometary particles. In: Gustafson BÅS, Hanner MS (eds) Physics, chemistry, and dynamics of interplanetary dust. Astronomical Society of the Pacific, San Francisco, pp 437–441

Xing Z, Hanner MS (1997) Light scattering by aggregate particles. Astron Astrophys 324:805–820

Xu Y-L (1995) Electromagnetic scattering by an aggregate of spheres. Appl Opt 34(21):4573–4588. doi:10.1364/AO.34.004573

Xu Y-L (2003) Radiative scattering properties of an ensemble of variously shaped small particles. Phys Rev E 67:046620. doi:10.1103/PhysRevE.67.046620

Xu Y-L, Gustafson BÅS (2001) A generalized multiparticle Mie-solution: further experimental verification. J Quant Spectrosc Radiat Transfer 70:395–419. doi:10.1016/S0022-4073(01)00019-X

Xu Y-L, Khlebtsov NG (2003) Orientation-averaged radiative properties of an arbitrary configuration of scatterers. J Quant Spectrosc Radiat Transfer 79–80:1121–1137. doi:10.1016/S0022-4073(02)00345-X

Yada T, Uesugi M, Uesugi K, Karouji Y, Suzuki Y, Takeuchi A, Tsuchiyama A, Okada T, Abe M (2014) Three dimensional structures of aggregate-type Itokawa particles. In: 77th annual meeting of the Meteoritical Society, p 5242

Yamamoto S, Kimura H, Zubko E, Kobayashi H, Wada K, Ishiguro M, Matsui T (2008) Comet 9P/Tempel 1: interpretation with the Deep Impact results. Astrophys J 673:L199–L202. doi:10.1086/527558

Yamamoto T, Hasegawa H (1977) Grain formation through nucleation process in astrophysical environment. Progress Theoret Phys 58:816–828. doi:10.1143/PTP.58.816

Yanamandra-Fisher PA, Hanner MS (1999) Optical properties of nonspherical particles of size comparable to the wavelength of light: application to comet dust. Icarus 138:107–128. doi:10.1006/icar.1998.6066

Yurkin MA, Hoekstra AG (2007) The discrete dipole approximation: an overview and recent developments. J Quant Spectrosc Radiat Transfer 106:558–589. doi:10.1016/j.jqsrt.2007.01.034

Yurkin MA, Hoekstra AG (2011) The discrete-dipole-approximation code ADDA: capabilities and known limitations. J Quant Spectrosc Radiat Transfer 112:2234–2247. doi:10.1016/j.jqsrt.2011.01.031

Zubko ES (2012) Light scattering by irregularly shaped particles with sizes comparable to the wavelength. In: Kokhanovsky AA (ed) Light scattering reviews 6. Springer, Berlin, pp 39–74. doi:10.1007/978-3-642-15531-4_2

Zubko E (2013) Light scattering by cometary dust: large-particle contribution. Earth Planet Space 65:139–148. doi:10.5047/eps.2012.02.003

Zubko E, Kimura H, Shkuratov Y, Muinonen K, Yamamoto T, Okamoto H, Videen G (2009) Effect of absorption on light scattering by agglomerated debris particles. J Quant Spectrosc Radiat Transfer 110:1741–1749. doi:10.1016/j.jqsrt.2008.12.006

Zubko E, Petrov D, Grynko Y, Shkuratov Y, Okamoto H, Muinonen K, Nousiainen T, Kimura H, Yamamoto T, Videen G (2010) Validity criteria of the discrete dipole approximation. Appl Opt 49(8):1267–1279. doi:10.1364/AO.49.001267

Zubko E, Furusho R, Kawabata K, Yamamoto T, Muinonen K, Videen G (2011) Interpretation of photo-polarimetric observations of comet 17P/Holmes. J Quant Spectrosc Radiat Transfer 112:1848–1863. doi:10.1016/j.jqsrt.2011.01.020

Zubko E, Muinonen K, Shkuratov Y, Hadamcik E, Levasseur-Regourd A-C, Videen G (2012) Evaluating the carbon depletion found by the Stardust mission in comet 81P/Wild 2. Astron Astrophys 544:L8. doi:10.1051/0004-6361/201218981

Zubko E, Muinonen K, Shkuratov Y, Videen G (2013) Characteristics of cometary dust in the innermost coma derived from polarimetry by Giotto. Mon Not R Astron Soc 430:1118–1124. doi:10.1093/mnras/sts679

Zubko E, Muinonen K, Videen G, Kiselev NN (2014) Dust in comet C/1975 V1 (West). Mon Not R Astron Soc 440:2928–2943. doi:10.1093/mnras/stu480

Part III
Polarimetry

Polarimetry of Man-Made Objects

Sergey N. Savenkov

Abstract Polarimetry has already been an active area of research for about fifty years. A primary motivation for research in scatter polarimetry is to understand the interaction of polarized radiation with natural scenes and to search for useful discriminants to classify targets at a distance. In order to study the polarization response of various targets, the matrix models (i.e., 2×2 coherent Jones and Sinclair and 4×4 average power density Mueller (Stokes) and Kennaugh matrices etc.) and coherent and incoherent target decomposition techniques has been used. This come to be the standard tools for targets characterization. Polarimetric decomposition methods allow a physical interpretation of the different scattering mechanisms inside a resolution cell. Thanks to such decompositions, it is possible to extract information related to the intrinsic physical and geometrical properties of the studied targets. This type of information is inestimable if intensity is measured only. The goal of the chapter is to explain the basics of polarimetric theory, outline its current state of the art, and review some of important applications to study the scattering behaviour of various man-made and urban targets like buildings (tall & short), ships, oil rigs and spills, mines, bridges etc. both in optical range and in radar polarimetry.

1 Introduction

To develop polarimetric methods for object identification and classification, one needs to understand the relation between polarimetric and physical properties of the objects. As electromagnetic radiation interacts with an object under study, its polarization state and intensity are changed. Polarization properties of the scattered radiation contain an ample of information on morphological and functional prop-

S.N. Savenkov (✉)
Faculty of Radio Physics, Electronics and Computer Systems,
Taras Shevchenko National University of Kyiv,
Volodymyrska Street, 64/13, Kiev 01601, Ukraine
e-mail: sns@univ.kiev.ua

© Springer-Verlag Berlin Heidelberg 2016

A. Kokhanovsky (ed.), *Light Scattering Reviews, Volume 11*,
Springer Praxis Books, DOI 10.1007/978-3-662-49538-4_9

erties of the object. For example, since depolarization of scattered radiation depends on the morphological and physical parameters of scatterers (i.e., density, size, distribution, shape, refractive index, etc.) present in the studied object (Elachi 1987; Boerner 1992; Bohren and Huffman 1983; Lee and Pottier 2009; Cloude 2010), this information can be utilized for making objects identification techniques. Many constituents of an object also exhibit polarization properties such as birefringence, dichroism, depolarization, etc., which might serve to discriminate between surface and volume scattering as well.

The enormous importance of the matrix optical and radar polarimetry is that it contains all the information that one can obtain from scattering scene (Bohren and Huffman 1983; Brosseau 1998; Azzam and Bashara 1987; Collett 1993; Shurcliff 1962). The matrix polarimetry has many useful applications in such diverse fields as interaction with various optical systems (Shurcliff 1962; Azzam and Bashara 1987; Collett 1993; Brosseau 1998), cloud diagnostics (van de Hulst 1957; Bohren and Huffman 1983; Mishchenko et al. 2000, 2002; Kokhanovsky 2003b), remote sensing of the ocean, atmosphere, and planetary surfaces (Boerner 1992; Kokhanovsky 2001, 2003a, b; Muttiah 2002; Mishchenko et al. 2010), biological tissue optics (Priezzhev et al. 1989; Tuchin 2002; Tuchin et al. 2006), and others.

The methods of interpretation of the Jones and Mueller matrices in optical polarimetry and target decompositions in radar polarimetry have been developed by many authors (Hurwitz and Jones 1941; Huynen 1970; Whitney 1971; Cloude 1986; Gil and Bernabeu 1986, 1987; Krogager 1990; Lu and Chipman 1994, 1996; Mar'enko and Savenkov 1994; Freeman and Durden 1998; Yamaguchi et al. 2005; Savenkov et al. 2005, 2006, 2007b).

In this chapter, we intend to illustrate the fact that polarization contributes reliably in wide scope of important applications to study the scattering behavior of various man-made and urban targets like buildings, oil contaminations, mines, bridges, etc., both in optical range and in radar polarimetry.

Many important polarimetric applications, which do not involve the matrix or full polarization measurement, are beyond the scope of our present discussion. In any case, our reference list should by no means be considered exhaustive and is merely intended to provide initial reference points for the interested reader.

2 Mueller Matrices of Deterministic and Depolarizing Objects

In the Mueller matrix calculus, the polarization state of light can be completely characterized by a Stokes vector, while the polarization transforming properties of a medium can be completely characterized by a Mueller matrix:

$$\mathbf{S}^{\text{out}} = \mathbf{M}\mathbf{S}^{\text{inp}}, \tag{1}$$

where the four-component Stokes column vector (with "out" and "inp" denoting the Stokes vectors of the output and input light, respectively) consists of the following parameters:

$$\mathbf{S} = \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} = \begin{pmatrix} \langle |E_x|^2 + |E_y|^2 \rangle \\ \langle |E_x|^2 - |E_y|^2 \rangle \\ \langle E_x^* E_y + E_x E_y^* \rangle \\ i\langle -E_x^* E_y + E_x E_y^* \rangle \end{pmatrix}, \tag{2}$$

with $i = (-1)^{1/2}$. Among the pioneering contributions to this field of research, we note those by Soleillet (1929), Perrin (1942), Mueller (1948), and Parke (1948, 1949).

The Stokes parameter I is proportional to the total energy flux of the light beam. The Stokes parameters Q and U represent the differences between two components of the flux in which the electric vectors oscillate in mutually orthogonal directions. The Stokes parameter V is the difference between two oppositely circularly polarized components of the flux. As indicated by the angular brackets, the Stokes parameters s_i are ensemble averages (or time averages in the case of ergodic, stationary processes). This implies that no coherence effects are considered.

The Stokes vectors and Mueller matrices represent operations on intensities and their differences, i.e., incoherent superpositions of light beams; they are not adequate to describe either interference or diffraction effects. However, they are well suited to describe partially polarized and unpolarized light. Extensive lists of various Mueller matrices have been presented by several authors (e.g., Shurcliff 1962; Kliger et al. 1990; Gerrard and Burch 1975).

The Stokes parameters obey the inequality

$$s_1^2 \geq s_2^2 + s_3^2 + s_4^2. \tag{3}$$

This inequality is called the Stokes–Verdet criterion and is a consequence of the Schwarz (or Cauchy-Bunyakovsky) theorem (Barakat 1963). The degree of polarization p is defined by

$$p = \sqrt{s_2^2 + s_3^2 + s_4^2} \Big/ s_1. \tag{4}$$

In Eq. (3), the equality holds for a completely polarized (pure) beam of light. In this case, $p = 1$. Another limiting case, $p = 0$, occurs when $s_2^2 + s_3^2 + s_4^2 = 0$, i.e., when the electric vector vibrates in all directions randomly and with no preferential orientation. An intermediate case, $0 < p < 1$, implies that light contains both polarized and depolarized components and is, therefore, called partially polarized.

The inequality (3) plays an important role in polarimetry because it allows one to classify the character of the light–medium interaction. Assume first that the input light is completely polarized. In this case, the equality in Eq. (3) implies that the medium is nondepolarizing. Note that the terms "nondepolarizing" and "deterministic" or "pure" are not, in general, identical. The term "deterministic" means that the Mueller matrix describing such a medium can be derived from the corresponding Jones matrix (Simon 1982; Gil and Bernabeu 1985; Anderson and Barakat 1994; Gopala Rao et al. 1998b). Hereinafter, we call this class of matrices pure Mueller matrices (Hovenier 1994). If the output light results in an inequality in Eq. (3) then the scattering medium is not deterministic. If, in addition, the transformation matrix in Eq. (1) can be represented as a convex sum of deterministic Mueller matrices (Cloude 1986; Gil 2000, 2007), then the result is a depolarizing Mueller matrix (hereinafter Mueller matrix); otherwise, the result is a Stokes transformation matrix, i.e., the transformation matrix ensures the fulfillment of the Stokes–Verdet criterion only. The properties of matrices transforming Stokes vectors into Stokes vectors, i.e., those satisfying the Stokes–Verdet criterion, have been studied by many authors (Xing 1992; van der Mee 1993; van der Mee and Hovenier 1992; Sridhar and Simon 1994; Nagirner 1993; Givens and Kostinski 1993; Gopala Rao et al. 1998a).

Any pure Mueller matrix $\mathbf{M}$ can be transformed into the corresponding Jones matrix $\mathbf{T}$ using the following relation (Parke 1949; Azzam and Bashara 1977; Dubois and Norikane 1987):

$$\mathbf{M} = \mathbf{A}(\mathbf{T} \otimes \mathbf{T}^*)\mathbf{A}^{-1}, \tag{5}$$

where the asterisk denotes the complex conjugate value,

$$\mathbf{T} = \begin{pmatrix} t_1 & t_4 \\ t_3 & t_2 \end{pmatrix}, \tag{6}$$

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & \mathrm{i} & -\mathrm{i} & 0 \end{pmatrix}, \tag{7}$$

the t_i are, in general complex, and $\otimes$ is the tensorial (Kronecker) product.

Since the element m_{11} is a gain for unpolarized incident light, it must satisfy the following inequality:

$$m_{11} > 0. \tag{8}$$

Furthermore, the elements of the Mueller matrix must obey the following conditions:

$$m_{11} \geq \left|m_{ij}\right|, \tag{9}$$

$$\mathrm{Tr}(\mathbf{M}) \geq 0, \tag{10}$$

$$\mu\mathbf{T} = |\mu|^2\mathbf{M}, \tag{11}$$

where Tr denotes the trace operation and μ is an arbitrary real or complex constant.

Note that the last relation defines the ability of the Jones and Mueller matrices to represent a "physically realizable" medium (Lu and Chipman 1994; Anderson and Barakat 1994; Gil 2007) and implies the physical restriction according to which the ratio g of the intensities of the emerging and incident light beams (the gain or intensity transmittance) must always be in the interval $0 \leq g \leq 1$. This condition is called the gain or transmittance condition and can be written in terms of the elements of the Mueller matrix as follows (Barakat 1987):

$$\begin{aligned} m_{11} + \left(m_{12}^2 + m_{13}^2 + m_{14}^2\right)^{1/2} &\leq 1, \\ m_{11} + \left(m_{21}^2 + m_{31}^2 + m_{41}^2\right)^{1/2} &\leq 1. \end{aligned} \tag{12}$$

While a Jones matrix has generally eight independent parameters, the absolute phase is lost in Eq. (5), yielding only seven independent elements for a pure Mueller matrix. Evidently, this results in the existence of interrelations for the elements of a general pure Mueller matrix. This fact was pointed out for the first time, although without a derivation of their explicit form, by van de Hulst (1957). Since then this subject has been studied by many authors (e.g., Abhyankar and Fymat 1969; Fry and Kattawar 1981; Hovenier et al. 1986). In the most complete and refined form these interrelations are presented in Hovenier (1994). In particular one can derive the following important equation for the elements of a pure Mueller matrix:

$$\sum_{i=1}^{4}\sum_{j=1}^{4} m_{ij}^2 = 4m_{11}^2. \tag{13}$$

This equality was obtained for the first time by Fry and Kattawar (1981). However, the question of whether this is a sufficient condition for $\mathbf{M}$ to be a pure Mueller matrix has been the subject of extensive discussions (see, e.g., Simon 1982, 1987; Hovenier 1994; Kim et al. 1987; Kostinski 1992; Kostinski et al. 1993; Gil and Bernabeu 1985; Anderson and Barakat 1994; Brosseau 1990; Brosseau et al. 1993). Under the premise that the Mueller matrix in question can be represented as a convex sum of pure Mueller matrices, Eq. (13) is both a necessary and a sufficient condition for $\mathbf{M}$ to be a pure Mueller matrix (Gil 2007).

In addition to the equalities presented above, a set of inequalities can be derived to characterize the structure of the pure Mueller matrix, as follows (Hovenier et al. 1986):

$$
\begin{aligned}
&m_{11}+m_{22}+m_{12}+m_{21}\geq 0,\\
&m_{11}-m_{22}-m_{12}+m_{21}\geq 0,\\
&m_{11}+m_{22}-m_{12}-m_{21}\geq 0,\\
&m_{11}-m_{22}+m_{12}-m_{21}\geq 0,\\
&m_{11}+m_{22}+m_{33}+m_{44}\geq 0,\\
&m_{11}+m_{22}-m_{33}-m_{44}\geq 0,\\
&m_{11}-m_{22}+m_{33}-m_{44}\geq 0,\\
&m_{11}-m_{22}-m_{33}+m_{44}\geq 0.
\end{aligned}
\tag{14}
$$

Equation (5) can be used to derive interrelations between the structures of a Jones matrix and the corresponding pure Mueller matrix. For example, the successive application of transposition and sign reversal for the off-diagonal elements of the Jones matrix in Eq. (6) yields

$$
\begin{pmatrix} t_1 & -t_3 \\ -t_4 & t_2 \end{pmatrix} \leftrightarrow \begin{pmatrix} m_{11} & m_{21} & -m_{31} & m_{41} \\ m_{12} & m_{22} & -m_{32} & m_{42} \\ -m_{13} & -m_{23} & m_{33} & -m_{43} \\ m_{14} & m_{24} & -m_{34} & m_{44} \end{pmatrix}. \tag{15}
$$

Physical reasons for the above relations are quite clear. Indeed, Eq. (15) originates from the operation of interchanging the incident and emerging light beams, the principle of reciprocity (Saxon 1955; Sekera 1966; Vansteenkiste et al. 1993; Potton 2004), and mirror symmetry (Hovenier 1969, 1970).

The effect of the symmetry of the individual scatterers and collections of scatterers on the structure (number of independent parameters) of the Mueller matrix has been considered by van de Hulst (1957). In particular, he demonstrated that the collection of scatterers containing equal number of particles and their mirror particles possesses the following Mueller matrix:

$$
\begin{pmatrix} m_{11} & m_{12} & 0 & 0 \\ m_{21} & m_{22} & 0 & 0 \\ 0 & 0 & m_{33} & m_{34} \\ 0 & 0 & m_{43} & m_{44} \end{pmatrix}. \tag{16}
$$

If in Eq. (6) $t_3 = t_4$, then $m_{11}^2 - m_{12}^2 - m_{33}^2 - m_{34}^2 = 0$, $m_{22} = m_{11}$, $m_{33} = m_{44}$, and $m_{34} = -m_{43}$, see Eq. (5).

If, in addition, the collection of scatterers contains equal number of particles in positions described by Eqs. (6) and (15) and those corresponding to the transposition and sign reversal of the off-diagonal elements of the Jones matrix (6), then $m_{12} = m_{21}$.

The Mueller matrix of Eq. (16) plays a key role in many light-scattering applications. Some of them will be discussed later in this section. The structure of the matrix Eq. (16) can be caused by a symmetry of individual particles and

a collection of particles in single and multiple scattering (van de Hulst 1957; Mishchenko and Travis 2000) as well as by "illumination-observation" geometries for backward and forward scattering (Zubko et al. 2004; Savenkov et al. 2007a).

The model of a medium described by the Mueller matrix of Eq. (16) has been used in studies of optical characteristics of oceanic water (Voss and Fry 1984; Kokhanovsky 2003a); ensembles of identical, but randomly oriented fractal particles (Kokhanovsky 2003c); dense spherical particle suspensions in the multiple-scattering regime (Kaplan et al. 2001); ice clouds consisting of nonspherical ice crystals in the multiple-scattering regime (Lawless et al. 2006); polydisperse, randomly oriented ice crystals modeled by finite circular cylinders with different size distributions (Xu et al. 2002); cylindrically shaped radially inhomogeneous particles (Manickavasagam and Menguc 1998). Other applications included measurements of the complex refractive index of isotropic materials as matrices of isotropic and ideal metal mirror reflections (Deibler and Smith 2001); the development of a symmetric three-term product decomposition of a Mueller–Jones matrix (Ossikovski 2008); and the description of very general and practically important cases of (i) randomly oriented particles with a plane of symmetry (Hovenier and van der Mee 2000) and/or (ii) equal number of particles and their mirror particles (Mishchenko et al. 2002). This list of applications can be extended significantly.

An example of the situation in which the Mueller matrix has the structure of Eq. (16) and contains information on the strong dependence of polarization and depolarization of output light on the polarization state of the input light is the exact forward scattering of polarized light by a slab of inhomogeneous linear birefringent medium (Savenkov et al. 2007a).

The scattering angles 0° (exact forward direction) and 180° (exact backward direction) deserve special attention owing to their importance in numerous practical applications, including the scattering by biological tissues. For the first time, the general form of Mueller matrices for these scattering angles was derived by van de Hulst (1957). Hu et al. (1987) presented a comprehensive study of forward and backward scattering by an individual particle in a fixed orientation. For forward scattering, they distinguished sixteen different symmetry shapes which were classified into five symmetry classes; for backward scattering, four different symmetry shapes were identified and classified into two symmetry classes. A large number of relations were derived in this way. The structures of Mueller matrices for various collections of particles in the cases of forward and backward scattering can be found elsewhere (van de Hulst 1957; Hovenier and Mackowski 1998).

It is important to note that although analyses of the internal structure of a general pure Mueller matrix, the symmetry relations between matrix elements caused by interchanging the incident and emerging light beams, and the principle of reciprocity have historically been carried out in the framework of light scattering by discrete particles; these results are also relevant to pure Mueller matrices in the continuous medium approximation.

In this section, we consider the problem of Mueller matrix interpretation in the framework of the approach wherein the medium studied is modeled as a medium

with a continuous (and possibly random) distribution of optical parameters. The polarization of light changes if the amplitudes and phases of the components of the electric vector **E** change separately or simultaneously (Shurcliff 1962; Azzam and Bashara 1987; Brosseau 1998). It is, therefore, customary to distinguish between the corresponding classes of anisotropic media: dichroic (or possessing amplitude anisotropy), influencing only the amplitudes; birefringent (or possessing phase anisotropy), influencing only the phases; and "all other" (possessing both amplitude and phase anisotropy) affecting both the amplitudes and the phases of the components of the electric field vector. Among these classes, four types of anisotropic mechanisms are recognized as basic or, after Jones, elementary (Jones 1941, 1942, 1947, 1956; Hurwitz and Jones 1941): linear and circular phase and linear and circular amplitude anisotropies.

Linear birefringence is described by the following pure Mueller matrix:

$$\mathbf{M}^{\mathrm{LP}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\alpha + \sin^2 2\alpha \cos\Delta & \cos 2\alpha \sin 2\alpha(1-\cos\Delta) & -\sin 2\alpha \sin\Delta \\ 0 & \cos 2\alpha \sin 2\alpha(1-\cos\Delta) & \sin^2 2\alpha + \cos^2 2\alpha \cos\Delta & \cos 2\alpha \sin\Delta \\ 0 & \sin 2\alpha \sin\Delta & -\cos 2\alpha \sin\Delta & \cos\Delta \end{pmatrix} \tag{17}$$

where Δ is the phase shift between two orthogonal linear components of the electric field vector and α is the azimuth of the anisotropy.

The Mueller matrix describing linear dichroism is

$$\mathbf{M}^{\mathrm{LA}} = \begin{pmatrix} 1+P & (1-P)\cos 2\gamma & (1-P)\sin 2\gamma & 0 \\ (1-P)\cos 2\gamma & \cos^2 2\gamma(1+P) + 2\sin^2 2\gamma\sqrt{P} & \cos 2\gamma \sin 2\gamma(1-\sqrt{P})^2 & 0 \\ (1-P)\sin 2\gamma & \cos 2\gamma \sin 2\gamma(1-\sqrt{P})^2 & \sin^2 2\gamma(1+P) + 2\cos^2 2\gamma\sqrt{P} & 0 \\ 0 & 0 & 0 & 2\sqrt{P} \end{pmatrix}, \tag{18}$$

where P is the relative absorption of two linear orthogonal components of the electric vector and $\gamma$is the azimuth of the anisotropy.

The Mueller matrix describing circular birefringence is

$$\mathbf{M}^{\mathrm{CP}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\varphi & \sin 2\varphi & 0 \\ 0 & -\sin 2\varphi & \cos 2\varphi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{19}$$

where φ is the induced phase shift between two orthogonal circular components of the electric vector.

At last, in terms of the Mueller matrix calculus, circular amplitude anisotropy is described by the following matrix:

$$\mathbf{M}^{\mathrm{CA}} = \begin{pmatrix} 1+R^2 & 0 & 0 & 2R \\ 0 & 1-R^2 & 0 & 0 \\ 0 & 0 & 1-R^2 & 0 \\ 2R & 0 & 0 & 1+R^2 \end{pmatrix}, \tag{20}$$

where R is the magnitude of anisotropy, i.e., the relative absorption of two orthogonal circular components of the electric vector. The six quantities α, Δ, P, γ, φ, and R are called anisotropy parameters.

It can be seen that the matrices describing linear and circular birefringence belong to the class of unitary matrices (in the case of matrices with real-valued elements—orthogonal matrices). The matrices of linear, Eq. (18), and circular, Eq. (20), dichroism belong to the class of Hermitian matrices (in the case of matrices with real-valued elements—symmetric matrices).

The Mueller matrices of Eqs. (17)–(20) represent media exhibiting individual types of anisotropy. Experimental measurements of these matrices or of the corresponding informative matrix elements allow one to interpret and characterize anisotropy properties of media. However, more often two or more types of anisotropy are exhibited by a medium simultaneously. Evidently, such cases require the development of more sophisticated polarimetric matrix models (Hurwitz and Jones 1941; Cloude 1986; Gil and Bernabeu 1987; Lu and Chipman 1996; Savenkov et al. 2006; Ossikovski 2008, 2009).

The matrix model that is extensively used in polarimetry for decoupling constituent polarization properties of optical medium is the polar decomposition proposed by Lu and Chipman (1996). This model is based on the so-called polar decomposition theorem (Lancaster and Tismenetsky 1985), according to which an arbitrary matrix $\mathbf{M}$ can be represented by a product

$$\mathbf{M} = \mathbf{M}_P\mathbf{M}_R \quad \text{or} \quad \mathbf{M} = \mathbf{M}_R\mathbf{M}'_P, \tag{21}$$

where $\mathbf{M}_P$ and $\mathbf{M}'_P$ are Hermitian matrices and $\mathbf{M}_R$ is a unitary one. The Hermitian matrix is associated with amplitude anisotropy, while the unitary matrix describes phase anisotropy (Whitney 1971). The matrices $\mathbf{M}_P$ and $\mathbf{M}_R$ are called the dichroic and the phase polar forms (Whitney 1971; Gil and Bernabeu 1987; Lu and Chipman 1996).

The polar decomposition was first employed by Whitney (1971) without finding explicit expressions for $\mathbf{M}_P$ and $\mathbf{M}_R$. They were proposed later, independently by Gil and Bernabeu (1987) and Lu and Chipman (1996). Alternatively, the dichroic and phase polar forms can be derived using spectral methods of linear algebra (Azzam and Bashara 1977).

The phase polar form $\mathbf{M}_R$ (using notation from Lu and Chipman 1996) is given by

$$\mathbf{M}_R = \begin{pmatrix} 1 & \vec{0}^T \\ \vec{0} & \mathbf{m}_R \end{pmatrix},$$
$$(\mathbf{m}_R)_{ij} = \delta_{ij}\cos R + a_i a_j(1-\cos R) + \sum_{k=1}^{3} \varepsilon_{ijk} a_k \sin R, \tag{22}$$

where $\vec{0}$ is the 3×1 zero vector; $[\,1 \quad a_1 \quad a_2 \quad a_3\,]^{\mathrm{T}} = [\,1 \quad \widehat{R}^{\mathrm{T}}\,]^{\mathrm{T}}$ is the normalized Stokes vector for the fast axis of $\mathbf{M}_R$; δ_{ij} is the Kronecker delta; ε_{ijk} is the Levi–Civita permutation symbol; $\mathbf{m}_R$ is the 3×3 submatrix of $\mathbf{M}_R$ obtained by striking out the first row and the first column of $\mathbf{M}_R$; and R is the birefringence given by

$$R = \arccos\left(\frac{1}{2}\mathrm{Tr}\,\mathbf{M}_R - 1\right). \tag{23}$$

The dichroic polar form $\mathbf{M}_P$ is as follows:

$$\mathbf{M}_P = T_u \begin{pmatrix} 1 & \vec{\mathbf{D}}^{\mathrm{T}} \\ \vec{\mathbf{D}} & \mathbf{m}_P \end{pmatrix}, \tag{24}$$

$$\mathbf{m}_P = \sqrt{1-D^2}\,\mathbf{I} + \left(1-\sqrt{1-D^2}\right)\widehat{\mathbf{D}}\widehat{\mathbf{D}}, \tag{25}$$

where $\mathbf{I}$ is the 3×3 identity matrix; $\widehat{\mathbf{D}} = \vec{\mathbf{D}}/\left|\vec{\mathbf{D}}\right|$ is the unit vector in the direction of the diattenuation vector $\vec{\mathbf{D}}$; T_{u} is the transmittance for unpolarized light; and the value of diattenuation can be obtained as

$$D = \left(1 - 4\left|\det(\mathbf{T})\right|^2 \Big/ \left[\mathrm{Tr}(\mathbf{T}^* \mathbf{T})\right]^2\right)^{1/2}. \tag{26}$$

The model of anisotropic media based on the polar decomposition contains in general six independent parameters, three for the phase polar form $\mathbf{M}_R$ and three for the dichroic polar form $\mathbf{M}_P$. It can be seen that the phase polar form is a unitary (orthogonal) matrix and the dichroic polar form is a Hermitian (symmetric) matrix. Note that unitarity (orthogonality) of the phase polar form, Eq. (22), is in complete agreement with the first Jones' equivalence theorem (Hurwitz and Jones 1941), and is a general model of elliptically birefringent media. The situation with the dichroic polar form is more complex (Savenkov et al. 2005, 2007b). Mathematically, the complexity originates from the fact that, in contrast to unitary matrices, the product of Hermitian matrices is generally not a Hermitian matrix (Lancaster and Tismenetsky 1985).

If the incident light is fully polarized and the output light is characterized by an inequality in Eq. (3) then the equalities for matrix elements obtained in Hovenier (1994) and Eq. (13), which determine the structure of the Mueller matrix as a deterministic matrix, are lost. In this case the output light is composed of several

incoherent contributions, and the medium as a whole cannot be represented by a Jones matrix. However, the medium can be considered as a parallel set of deterministic media, each one being described by a well-defined Jones matrix, in such a way that the light beam is shared among these different media. It is important to point out that the same result could be obtained by considering the medium as an ensemble (Kim et al. 1987) so that each realization i, characterized by a well-defined Jones matrix $\mathbf{T}^i$, occurs with a probability p^i.

If a Mueller matrix can be represented by a convex sum of pure Mueller matrices (Cloude 1986; Simon 1987; Cloude and Pottier 1995; Gil 2007), then it is called a depolarizing Mueller matrix. It is important to note that this class of matrices does not coincide with the class of matrices, called Stokes matrices, satisfying the Stokes–Verdet criterion, i.e., matrices transforming Stokes vectors into Stokes vectors, see Eq. (3). Any physical Mueller matrix is a Stokes matrix, but the converse is not, in general, true (Gil 2007). On the other hand, no method has been quoted to physically realize a Stokes matrix that cannot be represented as a convex sum of deterministic Mueller matrices.

In this case the following quadratic inequalities are also valid (Fry and Kattawar 1981):

$$\begin{aligned}
(m_{11}+m_{12})^2-(m_{21}+m_{22})^2 &\geq (m_{31}+m_{32})^2+(m_{41}+m_{42})^2,\\
(m_{11}-m_{12})^2-(m_{21}-m_{22})^2 &\geq (m_{31}-m_{32})^2+(m_{41}-m_{42})^2,\\
(m_{11}+m_{21})^2-(m_{12}+m_{22})^2 &\geq (m_{13}+m_{23})^2+(m_{14}+m_{24})^2,\\
(m_{11}-m_{21})^2-(m_{12}-m_{22})^2 &\geq (m_{13}-m_{23})^2+(m_{14}-m_{24})^2,\\
(m_{11}+m_{22})^2-(m_{12}+m_{21})^2 &\geq (m_{33}+m_{44})^2+(m_{34}-m_{43})^2,\\
(m_{11}-m_{22})^2-(m_{12}-m_{21})^2 &\geq (m_{33}-m_{44})^2+(m_{34}+m_{43})^2,
\end{aligned} \tag{27}$$

and Eq. (13) becomes an inequality as well:

$$\sum_{i=1}^{4}\sum_{j=1}^{4} m_{ij}^2 \leq 4m_{11}^2. \tag{28}$$

The study and characterization of depolarization is of considerable importance owing to the fact that depolarization phenomena are encountered in many theoretical and experimental applications of polarimetry to discrete random media and media with bulk and surface inhomogeneities. Note that the light–medium interaction with depolarization is heretofore studied in considerably less detail than the problem described by Mueller–Jones matrices discussed above.

Depolarization is the result of decorrelation of the phases and the amplitudes of the electric field vectors and/or selective absorption of polarization states (Brosseau 1998). Depolarization can be observed in both single and multiple light scattering and depends on geometrical and physical characteristics of the scatterers: shape, morphology, refractive index, size parameter (ratio of the particle circumference to

the wavelength of the incident light), and orientation with respect to the reference frame (Mishchenko and Travis 2000). Furthermore, multiple scattering results in depolarization of the output light even in the case of a collection of spherically symmetric particles and often reinforces depolarization caused by particle non-sphericity Mishchenko and Travis 2000; Mishchenko et al. 2006). Our purpose here is to characterize the depolarization phenomenon using the Mueller matrix formalism. In particular, we intend to discuss single-number depolarization metrics and Mueller matrices of depolarization.

Depolarization metric is a single scalar number that varies from zero, thereby corresponding to a totally depolarized output light, to a certain positive number corresponding to a totally polarized output light. All intermediate values are associated with partial polarization.

The depolarization index was introduced by Gil and Bernabeu (1985, 1986)

$$\mathrm{DI}(\mathbf{M}) = \sqrt{\sum_{i,j=1}^{4} m_{ij}^2 - m_{11}^2} \Big/ \left(\sqrt{3} m_{11}\right). \tag{29}$$

The depolarization index is bounded according to $0 \le \mathrm{DI}(\mathbf{M}) \le 1$. The extreme values of $\mathrm{DI}(\mathbf{M})$ correspond to the case of unpolarized and totally polarized output light, respectively.

An "analog" to the degree of polarization, Eq. (4), for linearly polarized input light in terms of Mueller matrix elements, the so-called index of linear polarization, was introduced by Bueno (2001):

$$G_{\mathrm{L}} = \frac{\sqrt{3}}{2m_{11}} \left(m_{21}^2 + m_{31}^2 + \frac{1}{3} \sum_{i=1}^{4} \left(m_{2i}^2 + m_{3i}^2 \right) \right)^{1/2}. \tag{30}$$

It can be seen that G_{L} is the ratio of the mean of the sum of the squares of matrix elements corresponding to linear polarization of the output light and the value of the corresponding averaged intensity normalized by the maximum value of this ratio which occurs for a linear polarizer: $(G_{\mathrm{L}})_{\max} = 2/3^{1/2}$. The former implies the following range of variation: $0 \le G_{\mathrm{L}} \le 1$.

The average degree of polarization was defined by Chipman (2005) as follows:

$$\text{Average DoP}(\mathbf{M}) = \frac{1}{4\pi} \int_{0}^{\pi} \int_{-\pi/2}^{\pi/2} p[\mathbf{MS}(\varepsilon, \zeta)] \cos\varepsilon \, \mathrm{d}\varepsilon \, \mathrm{d}\zeta. \tag{31}$$

The term $\cos\varepsilon \, \mathrm{d}\varepsilon \, \mathrm{d}\zeta$ scans the incident polarization state over the Poincaré sphere, with the latitude ε and longitude ζ. The Stokes vector $\mathbf{S}(\varepsilon, \zeta)$ is a function of ellipticity and orientation azimuth of the polarization ellipse of light:

$$\mathbf{S}(\varepsilon,\zeta) = [\,1 \quad \cos 2\varepsilon \cos 2\zeta \quad \cos 2\varepsilon \sin 2\zeta \quad \sin 2\varepsilon\,]^{\mathrm{T}}, \tag{32}$$

where T stands for "transposed."

The so-called $Q(\mathbf{M})$ metric is defined as follows (Espinosa-Luna and Bernabeu 2007):

$$Q(\mathbf{M}) = \left(3[\mathrm{DI}(\mathbf{M})]^2 - [D(\mathbf{M})]^2\right) \Big/ \left(1 + [D(\mathbf{M})]^2\right), \tag{33}$$

where $D(\mathbf{M}) = \left(m_{12}^2 + m_{13}^2 + m_{14}^2\right)^{1/2}$ is the diattenuation parameter and $0 \le D(\mathbf{M}) \le 1$. The metric $Q(\mathbf{M})$ is bounded according to $0 \le Q(M) \le 3$. Specifically, $Q(\mathbf{M}) = 0$ corresponds to a totally depolarizing medium; $0 < Q(M) < 1$ describes a partially depolarizing medium; $1 \le Q(M) < 3$ represents a partially depolarizing medium if, in addition, $0 < \mathrm{DI}(M) < 1$; otherwise, it represents a nondepolarizing diattenuating medium; finally, $Q(M) = 3$ for a nondepolarizing, nondiattenuating medium.

Thus, the depolarization metrics provide a summary of the depolarizing property of a medium via a single number. The depolarization index $\mathrm{DI}(\mathbf{M})$ and the $Q(\mathbf{M})$ metrics are directly related to the Mueller matrix elements only and, in contrast to the average degree of polarization (Average DoP), require no scan of the whole Poincaré sphere of the input polarizations. Potentially, $Q(\mathbf{M})$ provides more detailed information about depolarization properties of a medium.

Quantities referring to the intrinsic depolarization properties of light have had wide applications in polarimetry as well. Examples are the linear, δ_{L}, and circular, δ_{C}, depolarization ratios are defined according to (Mishchenko and Hovenier 1995; Mishchenko and Travis 2000)

$$\delta_{\mathrm{L}} = (s_1 - s_2)/(s_1 + s_2), \tag{34}$$

$$\delta_{\mathrm{C}} = (s_1 + s_4)/(s_1 - s_4). \tag{35}$$

The interest in these parameters is explained by the fact that they are susceptible to particle nonsphericity. Indeed, for spherical particles both ratios are equal to zero identically, whereas for nonspherical scatterers both δ_{L} and δ_{C} can substantially deviate from zero (Mishchenko and Hovenier 1995). The former means that if the incident light is linearly polarized then the backscattered light is completely linearly polarized in the same plane, whereas if the incident light is circularly polarized then the backscattered light is completely circularly polarized in the opposite sense. For nonspherical particles this is generally not the case.

Chipman (1995) introduced somewhat different versions of the degrees of linear and circular polarizations:

$$\mathrm{DoLP} = \sqrt{s_2^2 + s_3^2} \Big/ s_1, \tag{36}$$

$$\mathrm{DoCP} = s_4 / s_1. \tag{37}$$

These parameters turn out to be very useful for applications in meteorology, astronomy, ophthalmology, optical fibers, etc. (e.g., Bueno 2001 and references therein).

Some media depolarize all polarization states equally. Other depolarizing media partially depolarize most polarization states but may not depolarize one or some incident states. Depolarization depends significantly on the polarization state of the input light in the multiple-scattering regime (Bicout et al. 1994; Rojas-Ochoa et al. 2004; Kim et al. 2006, and references therein). In particular, Bicout et al. (1994) studied numerically and experimentally how depolarization evolves for linear and circular input polarizations as the size of the particles increases from very small (Rayleigh regime) to large (Mie regime) in the case of a forward scattering geometry.

A single-number metric providing a summary of depolarization by a medium cannot apparently give detailed information about all features of depolarization. Such information can only be obtained from Mueller matrix models of depolarization. The case when for all polarizations of the input light the degree of polarization p of the output light is the same which is called isotropic depolarization. When the degree of polarization of the output light is a function of parameters of the input polarization, one speaks of anisotropic depolarization.

There seems to be a consensus regarding the form of the Mueller matrix model describing isotropic depolarization (Brosseau 1998; Chipman 1999):

$$\mathrm{diag}[\,1 \quad p \quad p \quad p\,]. \tag{38}$$

It can be seen that the properties of this type of depolarization are the following:

(i) the transmittance is the same for all polarizations of the incident light;
(ii) p of the output light is the same for all input polarizations.

On the other hand, there is no consensus in the literature concerning the Mueller matrix for the case of anisotropic depolarization. Apparently, one of the most accepted forms of the Mueller matrix describing the dependence of p of the output light on the incident polarization is the following (Shindo 1995; Brosseau 1998; Chipman 1999; Ossikovski 2009):

$$\mathrm{diag}[\,1 \quad a \quad b \quad c\,]. \tag{39}$$

The elements a, b, and c are interpreted physically in the following manner: a and b are the degrees of linear depolarization, while c is the degree of circular depolarization. If $a = b = c = 0$ then the Mueller matrix represents an ideal depolarizer. Bicout et al. (1994) discussed the depolarization arising in multiple scattering

of light by spherical scatterers in the Rayleigh regime and is given by Eq. (39) with $a = b$.

The most general expression for the Mueller matrix describing depolarization was suggested by Lu and Chipman (1996):

$$\begin{pmatrix} 1 & \vec{\mathbf{O}}^{\mathrm{T}} \\ \vec{\mathbf{P}}_{\Delta} & \mathbf{m}_{\Delta} \end{pmatrix}, \quad \mathbf{m}_{\Delta}^{\mathrm{T}} = \mathbf{m}_{\Delta}, \tag{40}$$

where $\vec{\mathbf{P}}_{\Delta}$ denotes the so-called polarizance vector. The polarizance vector describes the state of polarization generated by this Mueller matrix from unpolarized incident light. The Mueller matrix of Eq. (40) has nine degrees of freedom, and this is of interest because this matrix along with a generalized deterministic Mueller matrix is jointly characterized by 16 degrees of freedom. This means that in this way one obtains the generalized Mueller matrix of an arbitrary medium that has 16 degrees of freedom and linearly interacts with polarized light.

The product of Mueller matrices of the unitary and Hermitian polar forms Eqs. (22) and (24) and the depolarizing Mueller matrix Eq. (40)

$$\mathbf{M} = \mathbf{M}_{\Delta}\,\mathbf{M}_{R}\,\mathbf{M}_{P}, \tag{41}$$

is the generalized polar decomposition and a multiplicative matrix model of an arbitrary Mueller matrix (Lu and Chipman 1996; Gil 2000, 2007).

The product of the phase polar form and the depolarizing matrices can then be obtained as

$$\mathbf{M}_{\Delta}\mathbf{M}_{R} = \mathbf{M}' = \mathbf{M}\mathbf{M}_{P}^{-1}. \tag{42}$$

Then

$$\vec{\mathbf{P}}_{\Delta} = (\vec{\mathbf{P}} - \mathbf{m}\vec{\mathbf{D}}) \Big/ (1 - D^2), \tag{43}$$

where $\vec{\mathbf{P}} = \big(1/m_{11}\big)[\, m_{21} \quad m_{31} \quad m_{41}\,]^{\mathrm{T}}$ and $\mathbf{m}$ is the submatrix of the initial matrix $\mathbf{M}$. The $\mathbf{m}'$ is the submatrix of $\mathbf{M}'$ and can be written as

$$\mathbf{m}' = \mathbf{m}_{\Delta}\mathbf{m}_{R}. \tag{44}$$

The submatrix $\mathbf{m}_{\Delta}$ can be calculated as follows:

$$\begin{aligned} \mathbf{m}_{\Delta} = \pm &\Big[\mathbf{m}'(\mathbf{m}')^{\mathrm{T}} + \Big(\sqrt{\lambda_1\lambda_2} + \sqrt{\lambda_2\lambda_3} + \sqrt{\lambda_1\lambda_3}\Big)\mathbf{I}\Big]^{-1} \\ &\times \Big[\Big(\sqrt{\lambda_1} + \sqrt{\lambda_2} + \sqrt{\lambda_3}\Big)\mathbf{m}'(\mathbf{m}')^{\mathrm{T}} + \sqrt{\lambda_1\lambda_2\lambda_3}\mathbf{I}\Big], \end{aligned} \tag{45}$$

where λ_i are the eigenvalues of $\mathbf{m}'(\mathbf{m}')^{\mathrm{T}}$. The sign "+" or "−" is determined by the sign of the determinant of $\mathbf{m}'$. The net depolarization coefficient Δ can be calculated according to

$$\Delta = 1 - \frac{1}{3}|\mathrm{Tr}(\mathbf{M}_\Delta) - 1|. \tag{46}$$

In conclusion of this section, we consider the additive Mueller matrix model of depolarizing object suggested by Cloude (1986) and extensively employed in optical and radar polarimetry (see, e.g., Savenkov et al. 2003; Savenkov and Muttiah 2004; Munoz et al. 2001, 2002, 2004; Volten et al. 2001; Cloude and Pottier 1996, 1997). The Cloude coherence matrix $\mathbf{J}$ is derived from the corresponding arbitrary Mueller matrix as follows:

$$\begin{aligned}
J_{11} &= (1/4)(m_{11} + m_{22} + m_{33} + m_{44}) & J_{12} &= (1/4)(m_{12} + m_{21} - im_{34} + im_{43}) \\
J_{13} &= (1/4)(m_{13} + m_{31} + im_{24} - im_{42}) & J_{14} &= (1/4)(m_{14} - im_{23} + im_{32} + m_{41}) \\
J_{21} &= (1/4)(m_{12} + m_{21} + im_{34} - im_{43}) & J_{22} &= (1/4)(m_{11} + m_{22} - m_{33} - m_{44}) \\
J_{23} &= (1/4)(im_{14} + m_{23} + m_{32} - im_{41}) & J_{24} &= (1/4)(-im_{13} + im_{31} + m_{24} + m_{42}) \\
J_{31} &= (1/4)(m_{13} + m_{31} - im_{24} + im_{42}) & J_{32} &= (1/4)(-im_{14} + m_{23} + m_{32} + im_{41}) \\
J_{33} &= (1/4)(m_{11} - m_{22} + m_{33} - m_{44}) & J_{34} &= (1/4)(im_{12} - im_{21} + m_{34} + m_{43}) \\
J_{41} &= (1/4)(m_{14} + im_{23} - im_{32} + m_{41}) & J_{42} &= (1/4)(im_{13} - im_{31} + m_{24} + m_{42}) \\
J_{43} &= (1/4)(-im_{12} + im_{21} + m_{34} + m_{43}) & J_{44} &= (1/4)(m_{11} - m_{22} - m_{33} + m_{44})
\end{aligned} \tag{47}$$

It can be seen that coherence matrix $\mathbf{J}$ is positive semidefinite Hermitian and, hence, has always four real eigenvalues. The eigenvalues of the coherence matrix, λ_i, can be combined to form a quantity that is a measure of the depolarization, depolarization metric, of the studied medium. This quantity is called entropy and is defined as

$$H = -\sum_{i=1}^{N} \left(\lambda_i \Big/ \sum\nolimits_j \lambda_j\right) \log_N \left(\lambda_i \Big/ \sum\nolimits_j \lambda_j\right) \tag{48}$$

Given eigenvalues λ_i of coherence matrix $\mathbf{J}$, we have for initial Mueller matrix:

$$\mathbf{M} = \sum_{k=1}^{4} \lambda_k \mathbf{M}_D^k; \quad \mathbf{M}_D^k \Leftrightarrow \mathbf{T}^k, \tag{49}$$

here $\mathbf{M}_D^k$ are the pure Mueller matrices obtained from the Jones matrices by Eq. (5).

The Jones matrix, $\mathbf{T}$, in turn, is obtained in the following manner:

$$\begin{aligned}
t_{11}^{(k)} &= \Psi_1^{(k)} + \Psi_2^{(k)}, \quad t_{12}^{(k)} = \Psi_3^{(k)} - i\Psi_4^{(k)} \\
t_{21}^{(k)} &= \Psi_3^{(k)} + i\Psi_4^{(k)}, \quad t_{22}^{(k)} = \Psi_1^{(k)} - \Psi_2^{(k)} \quad k = \overline{1,4}
\end{aligned}, \tag{50}$$

where $\boldsymbol{\Psi}^{(k)} = (\,\Psi_1 \quad \Psi_2 \quad \Psi_3 \quad \Psi_4\,)_k^T$ is kth eigenvector of coherence matrix $\mathbf{J}$.

Thus, the substance of the Cloude's coherency matrix concept, which, in essence, is an additive matrix model of depolarizing Mueller matrix, Eq. (49), is the representation of the initial depolarizing Mueller matrix as a weighted convex sum of four pure Mueller matrices.

If three of the eigenvalues of $\mathbf{J}$ vanish, then the entropy $H = 0$ and initial matrix $\mathbf{M}$ is a deterministic Mueller–Jones matrix. If all four eigenvalues of $\mathbf{J}$ are not equal to zero and $H \leq 0.5$, then the pure Mueller matrix, which corresponds to the maximal eigenvalue, is the dominant type of deterministic polarization transformation of the studied object. So, this model gives the possibility to study the anisotropy properties of depolarizing objects on the one hand and, on the other hand, it is a necessary and sufficient criterion for given 4×4 real matrix to be Mueller matrix (the case when all four eigenvalues of $\mathbf{J}$ are nonnegative) and pure Mueller matrix (the case when three of the eigenvalues vanish) (Munoz et al. 2001, 2002, 2004; Volten et al. 2001).

3 Mueller Matrix Polarimetry

The aim of this section is to discuss the general concept of the Mueller matrix measurement. An ample of practical schemes of the Stokes and Mueller polarimeters can be found elsewhere (Hauge 1980; Azzam 1997; Chipman 1995; Tyo et al. 2006).

The Mueller matrix polarimeter at visible and infrared is composed of a polarization state generator (PSG) and polarization state analyzer (PSA), as shown in Fig. 1.

The PSG forms the particular polarization state of incident light on the studied object. The PSA is operated to measure either the full Stokes vector or some of the Stokes parameters of the scattered light. Both PSG and PSA consist of retarders and diattenuators that are capable of analyzing the polarization state of the scattered beam.

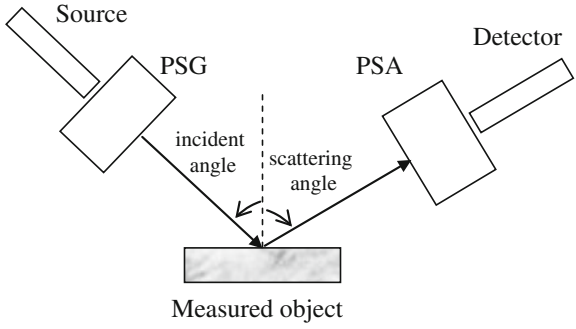


Fig. 1 Schematic overview to measure the Mueller matrices

Nearly all existing Mueller matrix polarimeters are configured so that the entire Mueller matrix has to be measured (Chipman 1995). Such approach is required to make up the well-conditioned set of 16 equations for matrix elements in order to reconstruct the full 4×4 Mueller matrix. The approach has been determined by the structure of the so-called characteristic or data reduction matrix of generalized measurement equation. The data reduction matrix describes the conversion of a set of polarized intensity measurements into the Mueller matrix which is represented as a 16×1 vector. For any PSG and PSA, the total flux measured by the detector is

$$g = \mathbf{QML} = \sum_{i=1}^{4} \sum_{j=1}^{4} q_i m_{ij} l_j, \tag{51}$$

where $\mathbf{L}$ is the Stokes vector produced by PSG; $\mathbf{M}$ is the object Mueller matrix; $\mathbf{Q}$ is the Stokes vector corresponding to the first row of the Mueller matrix representing the PSA.

To measure the full Mueller matrix, $N \geq 16$ flux measurements Eq. (51) are required. Flattening the Mueller matrix $\mathbf{M}$ into 16×1 Mueller vector of the form $\vec{\mathbf{M}} = [\, m_{11} \quad m_{12} \quad m_{13} \quad m_{14} \quad \cdots \quad m_{43} \quad m_{44} \,]^{\mathrm{T}}$ the polarimetric measurement equation can be represented as follows:

$$\mathbf{G} = \mathbf{W}\vec{\mathbf{M}} = \begin{pmatrix} q_1^1 l_1^1 & q_1^1 l_2^1 & q_1^1 l_3^1 & \cdots & q_4^1 l_4^1 \\ q_1^2 l_1^2 & q_1^2 l_2^2 & q_1^2 l_3^2 & \cdots & q_4^2 l_4^2 \\ q_1^3 l_1^3 & q_1^3 l_1^3 & q_1^3 l_3^3 & \cdots & q_4^3 l_4^3 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ q_1^N l_1^N & q_1^N l_2^N & q_1^N l_3^N & \cdots & q_4^N l_4^N \end{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ \cdots \\ m_{44} \end{pmatrix}, \tag{52}$$

where $\mathbf{G}$ is $N \times 1$ vector, whose components are the fluxes measured by detector; $\mathbf{W}$ is $N \times 16$ general characteristic or data reduction matrix with elements $w_{ij}^N = q_i^N l_j^N$.

Equation (52) is a system of generally N algebraic equations for Mueller matrix elements m_{ij}. The simplest case of the system Eq. (52) occurs when 16 independent measurements are performed. In this case $N = 16$, $\mathbf{W}$ is of rank 16, and inverse matrix $\mathbf{W}^{-1}$ is unique. Then all 16 Mueller matrix elements are

$$\vec{\mathbf{M}} = \mathbf{W}^{-1}\mathbf{G}. \tag{53}$$

Most Mueller matrix polarimeters are configured so that $N > 16$. This makes $\vec{\mathbf{M}}$ overdetermined, and $\mathbf{W}^{-1}$ does not exist. The optimal (least squares) estimation of $\vec{\mathbf{M}}$ can be obtained using the pseudoinverse matrix $\widetilde{\mathbf{W}}$ of $\mathbf{W}$ (Lankaster and Tismenetsky 1985):

$$\vec{\mathbf{M}} = \widetilde{\mathbf{W}}\mathbf{G} = \left(\mathbf{W}^{\mathrm{T}}\mathbf{W}\right)^{-1}\mathbf{W}^{\mathrm{T}}\mathbf{G}. \tag{54}$$

In mathematics, there exist a variety of pseudoinverse matrices (e.g., one-sided inverse, Drazin inverse, group inverse, Bott–Duffin inverse, etc.). Here, we use the so-called Moore–Penrose pseudoinverse matrix (Moore 1920; Bjerhammar 1951; Penrose 1955). Note that the characteristic matrix $\mathbf{W}$ in Eq. (52) is or can evidently be reduced to those of the full column rank.

This approach is named as the complete Mueller polarimetry. The theory of operation and calibration of Mueller matrix polarimetry was developed in (Chipman 1995), and the general formalism has been applied by many authors to the optimization of Mueller matrix polarimeters in the presence of noise and measurement error (Savenkov 2002; Smith 2002; De Martino et al. 2003; Twietmeyer and Chipman 2008). This procedure is repeated at different scattering angles in order to determine the angular profile of the Mueller matrix.

However, in many applications reconstruction of the full Mueller matrix is not necessary (Savenkov 2002, 2007; Tyo et al. 2010; Oberemok and Savenkov 2003). First of all, some subset of matrix elements might completely describe scattering which is of interest and hence, these subsets can be considered as initial information for the solution of corresponding classes of inverse problems. Another reason making the measurement of complete Mueller matrix unnecessary is matrix symmetry. Illustrative example is the pure Mueller matrix with symmetry determined by the first Jones equivalence theorem (Hurwitz and Jones 1941; Hovenier 1994). This matrix is widespread in the literature (Tang and Kwok 2001; Swami et al. 2006) because it describes linear crystal optics without absorption. This approach is termed incomplete or partial Mueller polarimetry (Savenkov 2007; Tyo et al. 2010).

The exact sets of matrix elements, i.e., structures of incomplete Mueller matrices, which can be measured in the framework of any of measurement strategies (time sequential, dynamic, etc.), are also determined by the structure of the data reduction matrix of generalized measurement equation, Eq. (52). This corresponds to the third case in Eq. (52) occurring when $N < 16$ and $\mathbf{W}$ is of rank less than 16. The optimal estimation of $\vec{\mathbf{M}}$ is again obtained using the pseudoinverse matrix $\widetilde{\mathbf{W}}$. However, in this case only 15 or fewer Mueller matrix elements can be determined from the system Eq. (54), i.e., polarimetry is "incomplete" or "partial."

4 Radar Polarimetry

In this section, we discuss the polarimetric metrics and targets' decompositions used for data interpretation in radar polarimetry. In so doing, we restricted ourselves mainly by those, which are used in the sections to follow. The comprehensive review of this subject can be found, for example, in excellent book by Lee and Pottier (2009).

Radar polarimetry is a special class of methods and equipments to carry out the polarization characterization at micro- and radiowave. Being active tool radar polarimetry is able to "see through" clouds and dusty conditions, to sense a target at any time of day or night, and, to some extent, "look under" vegetation and soil depending on wave bands, surface characteristics (e.g., ice, desert sand, close canopy, etc.), and conditions (e.g., moisture content).

Monostatic radars use the same antenna to transmit and receive electromagnetic radiation and are limited to measurements at the exact backscattering direction. Bistatic radars use one or more additional receiving antennas which provide supplementary polarization information.

Radar polarimetry has different polarization configurations. The most widely used are the linear polarizations indicated as *HH*, *VV*, *HV,* and *VH* where the first term refers to the polarization of input (emitted) and the latter to the received radiations, respectively. In this sense, the radar polarimetry can have different polarization levels:

- single polarization—*HH* or *VV* or *HV* or *VH*;
- dual polarizations—*HH* and *HV*, *VV* and *VH*, or *HH* and *VV*;
- four (quad) polarizations—*HH*, *VV*, *HV,* and *VH*.

The modes *HV* or *VH* are termed cross-polarization, whereas *HH* and *VV* modes are denoted as standard polarization. Quad-polarization polarimetry provides the four polarizations as *HH*, *VV*, *HV,* and *VH*, and also measures the difference in the magnitudes and phase between channels. Fully polarimetric sensors provide data that can be created using all possible combinations of transmitting and receiving polarizations, not the standard *HH* and *VV* only (Evans et al. 1988; Boerner et al. 1998).

The microwave electromagnetic ranges, in which radar polarimetry operates, are traditionally denoted by the letters shown in Table 1.

Table 1 Radar polarimetry bands and frequencies (after Lasaponara and Masini 2013)

Name	Nominal frequency range	Wavelength range	Specific bands used in SARs
VHF	30–300 MHz	10–1 m	138–144 MHz, 216–225 MHz
P (UHF)	300–1000 MHz	100–30 cm	420–450 MHz, 890–942 MHz
L	1–2 GHz	30–15 cm	1.215–1.4 GHz
S	2–4 GHz	17–7.5 cm	2.3–2.5 GHz, 2.7–3.7 GHz
C	4–8 GHz	7.5–3.75 cm	5.25–5.925 GHz
X	8–12 GHz	3.75–2.5 cm	8.5–10–68 GHz
Ku	12–18 GHz	2.5–1.67 cm	13.4–14.0 GHz, 15.7–17.7 GHz
K	18–27 GHz	1.67–1.11 cm	24.05–24.25 GHz
Ka	27–40 GHz	1.11–0.75 cm	33.4–36.0 GHz
V	40–75 GHz	0.75–0.40 cm	59–64 GHz
W	75–110 GHz	0.40–0.27 cm	76–81 GHz 92–100 GHz
Millimeter	110–300 GHz	2.7–1.0 mm	

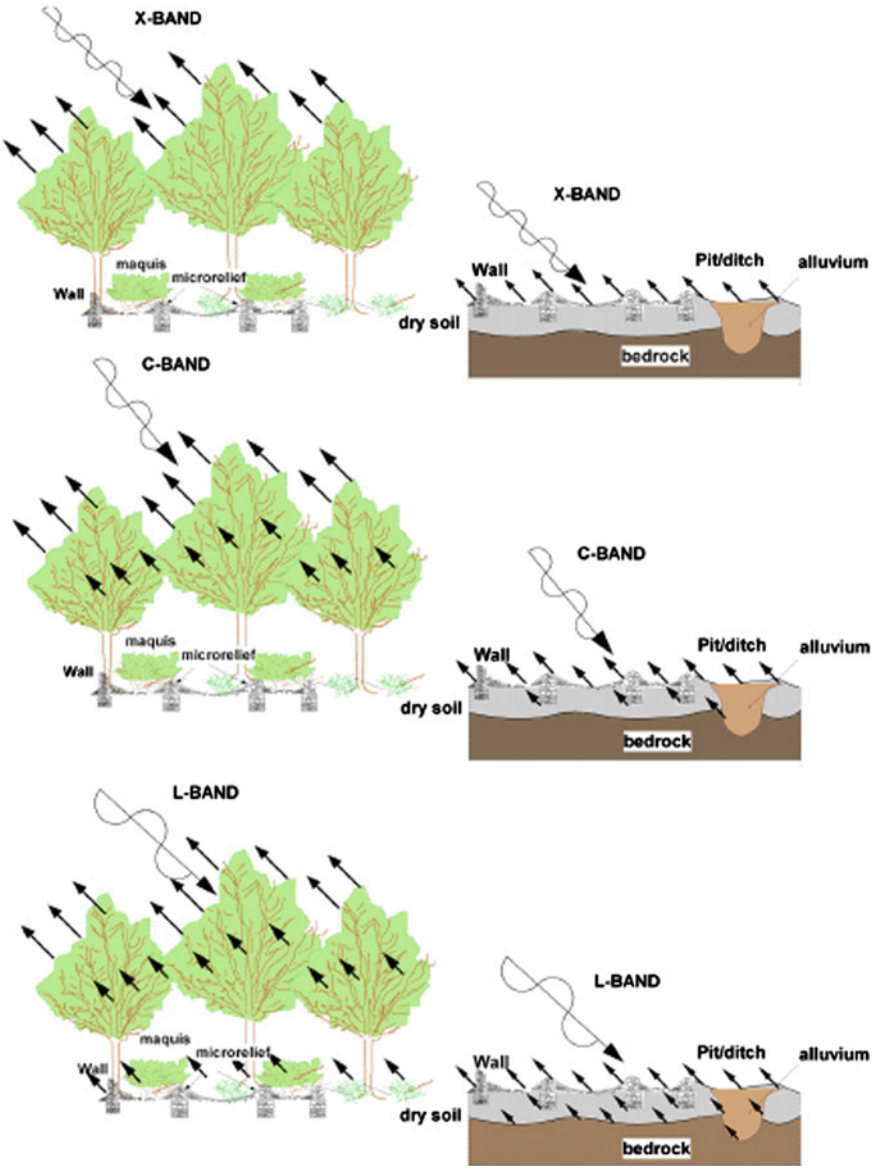


Fig. 2 Different penetration capabilities of radar polarimetry according to bands, land cover, and surface characteristics (after Lasaponara and Masini 2013)

Different frequencies are characterized by different "penetration capabilities" as schematically shown in Fig. 2.

SARs can have different polarization configurations. The most commonly used polarimetric SAR systems are presented in Table 2.

Table 2 SAR system parameters (after Lasaponara and Masini 2013)

SAR system	Band	Polarization	Incident angle (°)	Resolution (mt)	Swath width (km)	Organization	Altitude (km)	Orbit inclination (°)	Launch year
SEA-SAT	L	*HH*	23	25	100	NASA	790	108	1978
SIR-A	L	*HH*	45	30	50	NASA	225	57	1981
SIR-B	L	*HH*	20–60	30	50	NASA	225	57	1984
ALMAZ-1	S	*HH*	30–60	15	20–45	RSA (PKA)	300	72.7	1991
ERS-1	C	*VV*	24	25	100	ESA	790	97.7	1991
JERS-1	L	*HH*	35	18	76	NSDA/MITI	568	97.7	1992
SIR-C	C,L	All	17–60	25	15–100	NASA	225	57	1994
X-SAR	X	*VV*	17–60	25	15–40	DLR/ASI	228	57	1994
ERS-2	C	*HH*	24	25	100	ESA	785	97.7	1995
SAR-SAT	C	*HH*	17–50	10–100	50–170	CSA	790	98.6	1995
PRIRODA	S,L	*HH*	35	30	120	RSA/DLR	394	51.6	1995
ENVISAT	C	All	20–45	30	50–400	ESA	800	100	1998
SRTM	C	*HH*	20–60	30	60	NIMA/NASA	233	57	2000
PALSAR	L	*HH*; *VV HV*; *VH*	20–55	10–100	70–250	NASDA/MITI	700	98	2002
Light SAR	L	All	20	25–100	50–500	NASA	790	97.7	2003
COSMO-SkyMed ScanSar	X	One and two polarization modes (*HH*, *VV*, *HV*, or VH)	20–59	16–100	100–200	ASI	620	97.8	2007

(continued)

Table 2 (continued)

SAR system	Band	Polarization	Incident angle (°)	Resolution (mt)	Swath width (km)	Organization	Altitude (km)	Orbit inclination (°)	Launch year
COSMO-SkyMed StripMap	X		20–59	3–20	30–40	ASI	620	97.8	2007
COSMO-SkyMed SpotLight-2	X		20–59	1	10	ASI	620	97.8	2007
Terra SAR-X StrimMap mode	X	(*HH,VV*)–(*HH/VV*, *HH/HV*, *VV/VH*)	15–60	1.55–3.21	15–30	DLR-ASTIRUM	514	97.44	2007
Terra SAR-X SpotLight mode	X	(*HH,VV*)–(*HH/VV*)	15–60	1.34–3.21	10	DLR-ASTIRUM	514	97.44	2007
Terra SAR-X ScanSar mode	X		15–60	1.55–3.21	100	DLR-ASTIRUM	514	97.44	2007

For radar polarimetry, the Stokes vector is not the most effective way to characterize the data since there are effectively two measurements of polarization to quantify—one for each of the orthogonal transmitted waves. The radar transmits a horizontal polarized wave, measures the echo polarization, transmits a vertical polarized wave, and measures the polarization of that echo. At least two Stokes vectors would then be required (Woodhouse 2006). Since the polarimetric measurements of the echoes are made as orthogonal measurements it is convenient to define an alternative notion of the Jones matrix, Eq. (6), named in radar polarimetry as scattering matrix (Mott 2007; Cloude 2010)

$$\mathbf{S} = \begin{pmatrix} s_{VV} & s_{VH} \\ s_{HV} & s_{HH} \end{pmatrix}, \tag{55}$$

where s_{HV} denotes a transmitting antenna of a horizontal polarization and a receiving antenna of a vertical polarization. If the target is reciprocal $s_{HV} = s_{VH}$ then $\mathbf{S}$ becomes symmetric and has only three independent elements. Having measured this matrix, the strength and polarization of the scattered radiation for an arbitrary polarization of the incident radiation can be determined.

In Eqs. (6) and (55) there are two major conventions currently used for coordinates systems, the forward scatter alignment (FSA) convention and the backscattered alignment (BSA) convention (van Zyl and Zebker 1990; Guissard 1994; Boerner et al. 1998). If BSA convention is used, then the scattering matrix defined in Eq. (18) relates the scattered wave viewed approaching the receiving antenna to the incident wave viewed receding from the transmitting antenna (van Zyl et al. 1987). In optical or transmission polarimetry, the FSA convention is used.

The $\mathbf{S}$ matrix, which is expressed in the BSA coordinates, is referred to as the Sinclair matrix (Sinclair 1950; Kennaugh 1952; Zebker and van Zyl 1991; Guissard 1994; Boerner et al. 1998).

When we say in Eq. (1) that the input and output Stokes vectors are connected through the Mueller matrix, it assumes utilization of the FSA convention. Using the BSA convention, the Stokes vector of the backscattered wave is related to the incident-wave Stokes vector through the Kennaugh matrix, $\mathbf{K}$ (Kennaugh 1951; van de Hulst 1957; van Zyl et al. 1987; Boerner et al. 1998). Thus, the Mueller matrix $\mathbf{M}$ and the Kennaugh matrix $\mathbf{K}$ are formally related by (Guissard 1994; Luneburg 1995; Boerner et al. 1998)

$$\mathbf{M} = \mathrm{diag}[\,1 \quad 1 \quad 1 \quad -1\,] \cdot \mathbf{K}. \tag{56}$$

For symmetrical targets the matrices $\mathbf{S}$, $\mathbf{K}$, and $\mathbf{M}$ are characterized by only five parameters.

In addition to $\mathbf{M}$ and $\mathbf{K}$, two matrices named the target covariance matrix and the target coherence matrix might be used for characterization of partially polarized waves.

The scattering matrix of Eq. (55) can be represented in the vector form as follows:

$$\overrightarrow{\mathbf{S}} = [\, s_{HH} \quad s_{HV} \quad s_{VH} \quad s_{VV} \,]^{\mathrm{T}}. \tag{57}$$

An ensemble average of the complex product between $\overrightarrow{\mathbf{S}}$ and $\overrightarrow{\mathbf{S}}^{*\mathrm{T}}$ leads to the so-called covariance matrix **C** (Borgeaud et al. 1987; van Zyl and Ulaby 1990):

$$\mathbf{C} = \left\langle \overrightarrow{\mathbf{S}} \cdot \overrightarrow{\mathbf{S}}^{*\mathrm{T}} \right\rangle. \tag{58}$$

The Hermitian positive semidefinite matrix **C** has precisely the same elements as the Kennaugh matrix **K** and the Mueller matrix **M** but with different arrangements. The full-established properties of Hermitian matrices make convenient the use of **C** in some applications (Borgeaud et al. 1987; Kong et al. 1987; Novak and Burl 1990; van Zyl 1992; Touzi and Lopes 1994).

The coherency matrix Eq. (47) introduced by Cloude (1986) can per analogy with Eq. (58) be written as

$$\mathbf{J} = \left\langle \overrightarrow{\mathbf{k}} \cdot \overrightarrow{\mathbf{k}}^{*\mathrm{T}} \right\rangle, \tag{59}$$

where $\overrightarrow{\mathbf{k}}$ is the target scattering vector (Cloude 1986) and is given by

$$\overrightarrow{\mathbf{k}} = [\, s_{HH} + s_{VV} \quad s_{HH} - s_{VV} \quad s_{HV} + s_{HV} \quad i(s_{HV} - s_{VH}) \,]^{T}. \tag{60}$$

Note that **C** and **J** are unitarily similar. The two matrices carry the same information; both are Hermitian positive semidefinite and both have the same eigenvalues but different eigenvectors (Cloude and Pottier 1996).

Once reciprocity, $s_{HV} = s_{VH}$, has been assumed that the elements of **S** may be stacked into a three-element vector $\overrightarrow{\mathbf{k}} = [\, s_{VV} \quad s_{HV} \quad s_{HH} \,]^{\mathrm{T}}$. The linear basis is not always the most efficient way of dealing with the analysis of polarimetric data, and the Pauli basis of the target vector

$$\overrightarrow{\mathbf{k}_P} = [\, s_{HH} - s_{VV} \quad 2s_{HV} \quad s_{HH} + s_{VV} \,]^{\mathrm{T}} \tag{61}$$

is for many applications more useful as it helps to emphasize the phase difference between the *HH* and *VV* terms. Double interactions are dominated by the first term, multiple (volume) scattering dominates the second term, and direct scattering is dominated by the second term.

In order to extract information from the polarized backscatter data and obtain physical interpretation of the Mueller matrix **M**, the Kennaugh matrix **K**, the target covariance matrix and coherence matrices, a number of polarimetric metrics and targets decompositions have been devised (see, e.g., Freeman and Durden 1998; Cloude and Pottier 1996, 1997). The target decomposition models attempt to characterize the backscatter as a sum of elementary scattering mechanisms such as single bounce, double bounce, and volume scattering, see Fig. 3.

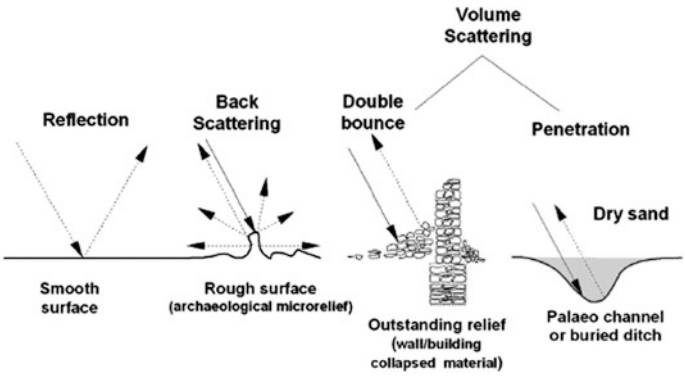


Fig. 3 Elementary scattering mechanisms, from *left* to *right*: single bounce (smooth surface), double bounce (urban settlement remains or upstanding relief), and volume scattering (urban settlement remains and soil penetration) (after Lasaponara and Masini 2013)

Equation (59) gives the 3×3 coherency matrix and corresponding three components' decomposition of the target Mueller matrix. Using normalized eigenvalues of coherency matrix $p_i = \lambda_i / \sum_{i=1}^{3} \lambda_i$ (Cloude et al. 2001) one more important polarimetric, the scattering anisotropy A, can be introduced,

$$A = (p_2 - p_3)/(p_2 + p_3). \tag{62}$$

Scattering anisotropy A varies between zero and one defining the relation between the second and the third eigenvalues, i.e., the difference of the secondary scattering mechanisms. Entropy for smooth surfaces becomes zero implying a nondepolarizing scattering process described by a single scattering matrix and increases with surface roughness. Depolarizing surfaces are characterized by nonzero entropy values. However, A can be zero even for rough surfaces. For surfaces characterized by intermediate entropy values, a high scattering anisotropy A indicates the presence of only one strong secondary scattering process, while a low anisotropy indicates the appearance of two equally strong scattering processes. For azimuthally symmetric surfaces $p_2 = p_3$ and A becomes zero (Cloude and Pottier 1996). In this sense, the anisotropy provides complementary information to the entropy and facilitates the interpretation of the surface scatterer.

At last, eigenvalue–eigenvector decomposition of the coherency matrix Eq. (59) provides one more metric, the alpha angle α, representing the dominant scattering mechanism and is calculated from the eigenvectors and eigenvalues of $\mathbf{J}$:

$$\alpha = \sum_{i=1}^{3} (\alpha_i p_i), \tag{63}$$

where α_i are the scattering mechanisms represented by the three eigenvectors: $\alpha = 0$ corresponds to a surface scattering, $\alpha = 45^\circ$ to a volume scattering, and $\alpha = 90^\circ$.

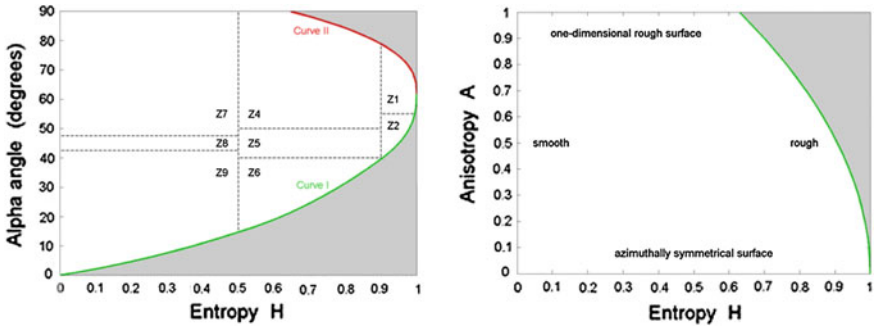


Fig. 4 The H/α and H/A planes (after Cloude and Pottier 1997)

In Cloude and Pottier (1997) the useful geometrical interpretation of the information in the coherency matrix based on $H/A/\alpha$ set of parameters in form of nine classes has been suggested, see Fig. 4.

The class interpretations suggested by Cloude and Pottier are as follows:

Z1 Double bounce scattering in a high entropy environment
Z2 Multiple scattering in a high entropy environment (e.g., forest canopy)
Z3 Surface scattering in a high entropy environment (not a feasible region in H/α space)
Z4 Medium entropy multiple scattering
Z5 Medium entropy dipole scattering
Z6 Medium entropy surface scattering

Because of the averaging of the different scattering mechanisms, i.e., averaging of the different eigenvectors, restricts the range of the possible values of α and the observable values of α for a given entropy are bounded between curves I and II as the entropy increases.

The H / A plane has been used to distinguish different types of surface scattering.

Note the parameters $H/A/\alpha$ are independent of rotation of the target about the radar line of sight. The latter in turn means that the parameters can be obtained independent on the polarization basis. In addition, the classes above are chosen based on general properties of the scattering mechanisms and do not depend on a particular dataset. This allows to carry out an unsupervised classification of the scenes under consideration according to the type of scattering process within the sample H, A, and the corresponding physical scattering mechanism α.

Another set of polarimetric metrics to characterize the scattering scene can be deduced from Pauly decomposition of the scattering matrix Eq. (55) (Lee and Pottier 2009). In accordance with this decomposition the scattering matrix **S** is represented as the complex sum of the Pauli matrices. The elementary scattering mechanisms are associated with corresponding basis matrices:

$$\mathbf{S} = \frac{a}{\sqrt{2}}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} + \frac{b}{\sqrt{2}}\begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} + \frac{c}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \frac{d}{\sqrt{2}}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad (64)$$

where a, b, c, and d are complex and given by

$$\begin{aligned} a &= (s_{HH} + s_{VV})/\sqrt{2}, \quad b = (s_{HH} - s_{VV})/\sqrt{2} \\ c &= (s_{HV} + s_{VH})/\sqrt{2}, \quad d = i(s_{HV} - s_{VH})/\sqrt{2} \end{aligned} \quad (65)$$

In the monostatic case, where $s_{HV} = s_{VH}$, the Pauli matrix basis can be reduced to the first three matrices yeilding $d = 0$.

The Pauli decomposition determines the following three scattering mechanisms characterizing the target under consideration: $|a|^2$ determines the power scattered by the targets characterized by single or odd bounce; $|b|^2$ determines the power characterized by double or even bounce; $|c|^2$ determines the power characterized by a deplane oriented at 45°. The scattering mechanism represented by $|c|^2$ refers to backscattering with orthogonal polarization. One of the character examples of this is the volume scattering produced by the forest canopy.

The scattering patterns are described by four independent variables, the ellipticity and orientation of the incident wave, and the ellipticity and orientation of the backscattered wave. However, using all four of these variables would result in a response that would be too cumbersome and complicated to interpret. Therefore, for clearness, only two variables are used at a time to interpret the scattering patterns—the ellipticity and orientation angle of the incident wave. This yields two signatures—the copolarization and cross-polarization signatures introduced in (van Zyl et al. 1987; Agrawal and Boerner 1989; McNairn et al. 2002), see Fig. 5. In the copolarization case, the polarization of the scattered wave is the same as the polarization

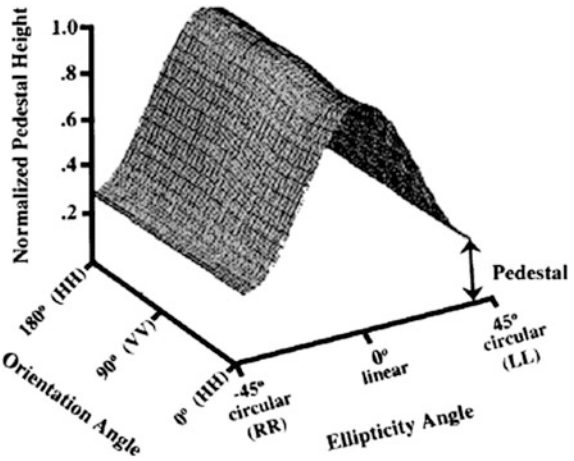


Fig. 5 Example of polarization signature (after McNairna et al. 2002)

of the incident wave, while in the cross-polarization case, the polarization of the scattered wave is orthogonal to the polarization of the incident wave.

The height of the pedestal, see Fig. 5, is an indicator of the presence of an unpolarized scattering component, and thus the degree of polarization of a scattered wave (van Zyl et al. 1987). Signatures with significant pedestals are typical of targets that are dominated by volume scattering or multiple surface scattering, e.g., Evans et al. (1988) reported that pedestal height was directly proportional to vegetation density. Ray et al. (1992) and van Zyl (1989) noted that pedestal height was related to surface roughness with increase in roughness resulting in higher pedestals.

Although being not unique, polarization signature capturing many scattering characteristics of the target, at all polarizations, is informative and can indicate the scattering mechanisms dominating the target response. Since different scattering mechanisms give different polarization signatures, they could be extracted from the measured Mueller matrix. The "building blocks" of such interpretation of the measured Mueller matrix are the following scattering mechanisms: (1) double bounce scattering; (2) Bragg scattering; (3) single (odd) bounce scattering; (4) cross-scattering (see, for example, Dong et al. 1996). Since measurements for independent scattering mechanisms can be added incoherently (van de Hulst 1957; Kim et al. 1987) then the total Mueller matrix can be represented as a sum of above basic scattering mechanisms. A number of additive decompositions of the scattering and Mueller matrix can be found elsewhere (Cloude and Pottier 1996, 1997; Touzi et al. 2004).

Double bounce scattering models are typically the scattering from the dihedral-corner-reflector-like structures such as the trunk-ground structure in forested areas and the wall-ground structures in urban areas. It has been shown (Dong and Richards 1995a) that the scattering matrix for this mechanism can be written as

$$\mathbf{S}_1 = \begin{pmatrix} 1 & 0 \\ 0 & (1/\sqrt{\alpha})\exp(i\delta) \end{pmatrix}, \tag{66}$$

where α and δ are referred to as polarization index defined as the ratio of *HH* to *VV* polarization response and polarization phase difference defined as the phase difference between *HH* and *VV* backscattered responses. For trunk-ground structure, $\alpha \approx 4-6$ and $\delta \approx 140^\circ - 160^\circ$ (Dong and Richards 1995b). The Mueller matrix for the double bounce scattering has the form

$$\mathbf{M}_1 = \begin{pmatrix} (\alpha+1)/2\alpha & (\alpha-1)/2\alpha & 0 & 0 \\ (\alpha-1)/2\alpha & (\alpha+1)/2\alpha & 0 & 0 \\ 0 & 0 & (1/\sqrt{\alpha})\cos\delta & -(1/\sqrt{\alpha})\sin\delta \\ 0 & 0 & -(1/\sqrt{\alpha})\sin\delta & -(1/\sqrt{\alpha})\cos\delta \end{pmatrix}. \tag{67}$$

Bragg scattering models the scattering from slightly rough surfaces, for example, sea surface (Valenzuela 1967; Elachi 1987). The Mueller matrix of the Bragg scattering is

$$\mathbf{M}_2 = \begin{pmatrix} (\alpha+1)/2\alpha & (\alpha-1)/2\alpha & 0 & 0 \\ (\alpha-1)/2\alpha & (\alpha+1)/2\alpha & 0 & 0 \\ 0 & 0 & (1/\sqrt{\alpha}) & 0 \\ 0 & 0 & 0 & -(1/\sqrt{\alpha}) \end{pmatrix}, \tag{68}$$

where the mean polarization index value $\alpha < 1$; the mean polarization phase difference value δ is zero, the Bragg scattering undergoes a single bounce.

The single bounce scattering models the direct specular reflections from the ground surfaces or from building roofs perpendicular to incident waves. The Mueller matrix for this mechanism is

$$\mathbf{M}_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{69}$$

The copolarized response from forest crown volume backscattering can be included in this mechanism. If the orientations of leaves, twigs, and small branches are assumed to be uniformly distributed, the backscattering response will be independent of polarization, giving the same *HH* and *VV* responses. The backscattering from the trihedral-corner-reflector-like, wall–wall–ground structures can also be classified in this mechanism, since the scattering undergoes odd bounces (Dong et al. 1996).

The polarimetric response of a point or distributed target can generally consist of the co- and cross-polarized responses. If the total cross-polarized component is of interest, it can be assumed theoretically that the total cross-polarized response is caused by a hypothetical cross-scattering mechanism whose scattering matrix is

$$\mathbf{S}_4 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{70}$$

and the corresponding Mueller matrix is

$$\mathbf{M}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{71}$$

5 Applications of Optical and Radar Polarimetry

Optical and radar polarimetry have recently demonstrated their unique abilities in identification, characterization, and classification of anisotropic media and point or distributed targets of various nature. Several examples of their applications will be discussed in this section because a comprehensive survey would here be impossible through, on the one hand, naturally restricted space of the chapter, and, on the other hand, all applications considered below actually deserve to date the individual books to be discussed systematically. Thus, we select only a few representative examples taken from different polarimetrical problems. Additionally, in all cases below we concentrate on polarimetric results literally. For more other computing, technical details one should refer to corresponding bibliography hereinafter.

5.1 Contamination

5.1.1 Ocean

Although, yet almost decade ago it was suggested to pioneer the use of the microwave polarimetric measurements for oil slick observation (Brown and Fingas 2003), first results appeared unsatisfactory (Gade et al. 1998) and not further developed. Only recently new studies showed the usefulness of fully and partially polarimetric SAR measurements (see, e.g., Migliaccio et al. 2007, 2009a; Nunziata et al. 2008; and references therein).

Sea oil spills and pollution is a matter of great concern since it affects ruinously and sometimes irretrievably both the environment and human health. Oil slick detection is determinative for effective planning of restorative countermeasures and to minimize pollution fatal consequences. In this sense, through its all-weather day and night capabilities (Brown and Fingas 2003), the full polarimetric radar data is undoubtedly recognized as the most informative and useful tool for global and regional environment monitoring, disaster management support, and resource survey.

The oil spill detection approach is based on the different sea surface scattering mechanisms expected with and without surface slicks (non-Bragg and Bragg, respectively) (Nunziata et al. 2008; Migliaccio et al. 2009a) and, therefore, they are characterized by different polarimetric features (see Table 3).

Table 3 Oil-free and oil-covered polarimetric features (after Migliaccio et al. 2009b)

	Oil-free	Oil-covered
Polarimetric entropy	Low	High
Unpolarized energy	Low	High
HH–VV correlation	High	Low

In (Migliaccio et al. 2009b) two approaches to the Mueller matrix analysis of three (one oil spill accident and two of look-alikes) data samples are presented. First, Mueller matrix analysis approach assumes using the copolarized signatures (see Sect. 4). Migliaccio et al. (2008) shows that a sea surface without slick, characterized by low unpolarized backscattered energy (see Table 3), yields a copolarized signature with a small pedestal height. The larger pedestal height corresponds to the case of a sea surface with slick, characterized by a large unpolarized backscattered signal. Second approach is based on the use of the Cloude entropy, Eq. (48). In scope of this approach in the case of a sea surface without slick low entropy is evidently expected, and high entropy is characteristic to the case of a sea surface with slick.

First data sample is oil spill accident occurred with tanker Solar I in August 11, 2006 sank in 640 m of water about 24 km off the southern coast of Lusaca Point, Guimaras Island of the Philippines, causing heavy oil spill (approximately 200,000 L) from the tanker.

In Fig. 6a polarimetric L band SAR data (*VV*) for accident area is presented. Figure 6b depicts this area in the polarimetric entropy terms. It can be seen that, as predicted in (Migliaccio et al. 2008), the entropy values for the oil-covered area are larger than the surrounding sea ones.

The normalized copolarized signature has been synthesized for both the oil-free, Fig. 7a, and the oil-covered, Fig. 7b, sea surfaces by considering two regions of interest of equal size. Again, comparison of the two polarization signatures shows that in accordance with (Migliaccio et al. 2008) the oil slick, increasing the amount of unpolarized backscattered energy yields the higher pedestal.

Two other cases relate to two look-alikes which belong to the acquisition on March 10, 2007, off the coasts of Da Nang (Vietnam). Figure 8a shows polarimetric L band SAR data (*VV*) for one of look-alike area in which a dark elongated feature is visible. Figure 8b is the same in the polarimetric entropy terms. In Fig. 9 the

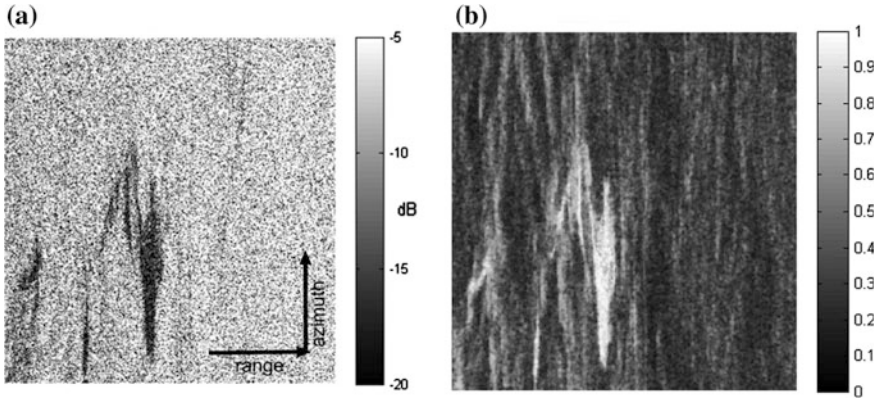


Fig. 6 **a** Polarimetric L band SAR data (*VV*) for accident area. **b** Polarimetric entropy (after Migliaccio et al. 2009b)

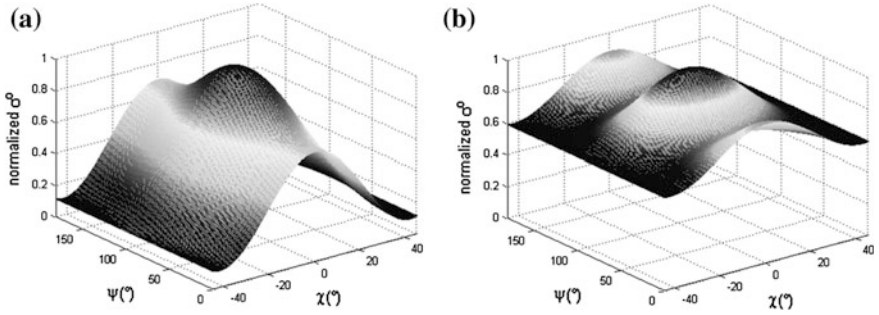


Fig. 7 The normalized copolarized signatures **a** oil-free and **b** oil-covered sea surfaces (after Migliaccio et al. 2009b)

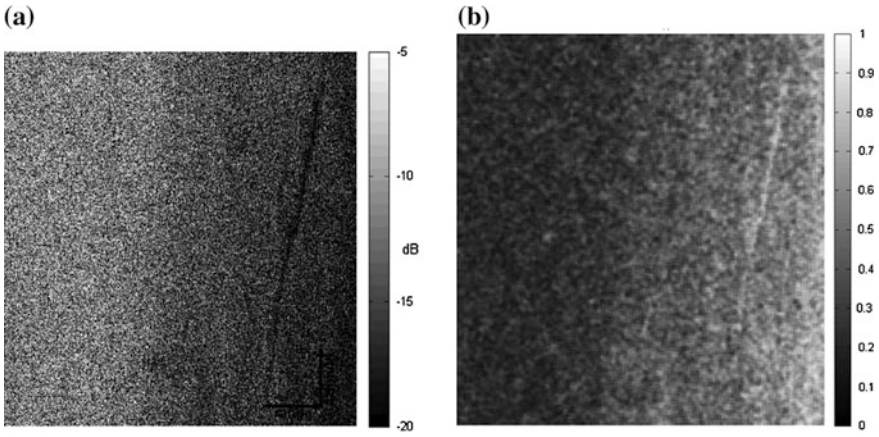


Fig. 8 **a** Polarimetric L band SAR data (*VV*) for look-alike area. **b** Polarimetric entropy (after Migliaccio et al. 2009b)

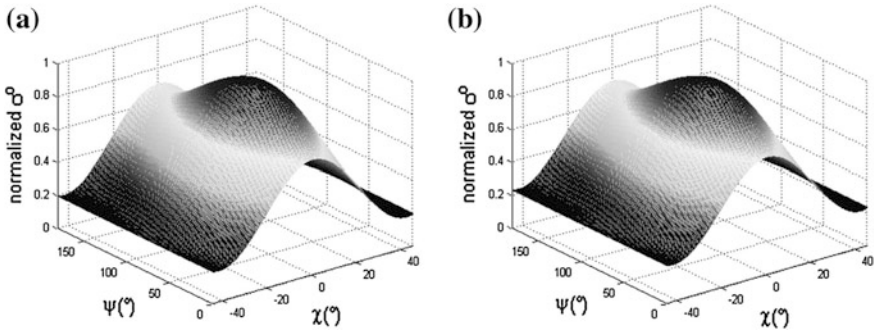


Fig. 9 Normalized copolarized signatures for **a** the slick-free sea surface and **b** the dark area (after Migliaccio et al. 2009b)

normalized copolarized signature, synthesized for both the slick-free sea surface (a) and the dark area (b) is presented.

The polarimetric entropy, Fig. 8b, does not show any reliable features which can be attributed to the dark area in Fig. 8a. Comparison of the two copolarized signatures, Fig. 30, shows low and almost equal pedestals for the slick-free sea surface and the dark area. These allow classifying the dark area in Fig. 8a as a look-alike.

The experimental results obtained show the effectiveness and the usefulness of the polarimetric L band SAR data for sea oil slick observation.

5.1.2 Chemical–Biological Materials

Undoubtedly, a need for the rapid and reliable identification of toxic chemical–biological materials and their derivatives in the atmosphere or on remote surfaces at safe distances (standoff detection) is urgent for military, industrial, and civilian defense practices (Demirev et al. 2005; Petryk 2007 and references therein). It is even more crucial because only long-range and beforehand detection and classification of toxic chemical–biological materials will allow one to take appropriate countermeasures minimizing their environment and population consequences.

In (Haugland et al. 1992; Carrieri 1999; Carrieri et al. 1998, 2007, 2008, and 2010) an approach to detect the presence of particular compounds in the atmosphere or on remote surfaces named differential absorption Mueller matrix spectroscopy (DIAMMS) has been suggested. The approach consists of twice the measurements of the Mueller matrix of the suspected contaminant in the backscattering direction, using two CO_2 laser emission lines as the light sources. The matter is that the Mueller matrices would not normally be measurably different for two close wavelengths. However, if one of the wavelengths is selected to coincide with a known IR absorption band of the suspected contaminant, a significant difference in one or more of the 16 Mueller matrix elements in switching between wavelengths can occur.

This approach was tested on number of materials, in particular, crystalline samples comprising amino acids, sugars, enantiomorphs, aerosolized γ-irradiated Bacillus subtilus and protein chicken egg white albumen, and some others. These materials were chosen because they are nontoxic, easy to handle, have some polarization properties that are inherent in the more complex biological organisms, and acting as biological warfare agent surrogates.

Samples of crystalline organic materials were prepared by pressing the powder to form the wafers, which could then be vertically arranged in polarimeter. The sample Mueller matrix for backscattering was measured for the orientation angle which is of normal incidence $\pm 20^{\circ}$. For aerosol measurements, materials were dispersed inside windowless pneumatic chambers, and then irradiated across $9.1-12.0$ μm optical bandwidth. Each aerosol was interrogated between 30 and 94 beam wavelengths, and at least one fundamental molecular vibration or vibration–rotation resonance mode in the analyte is stimulated from these infrared probing beams.

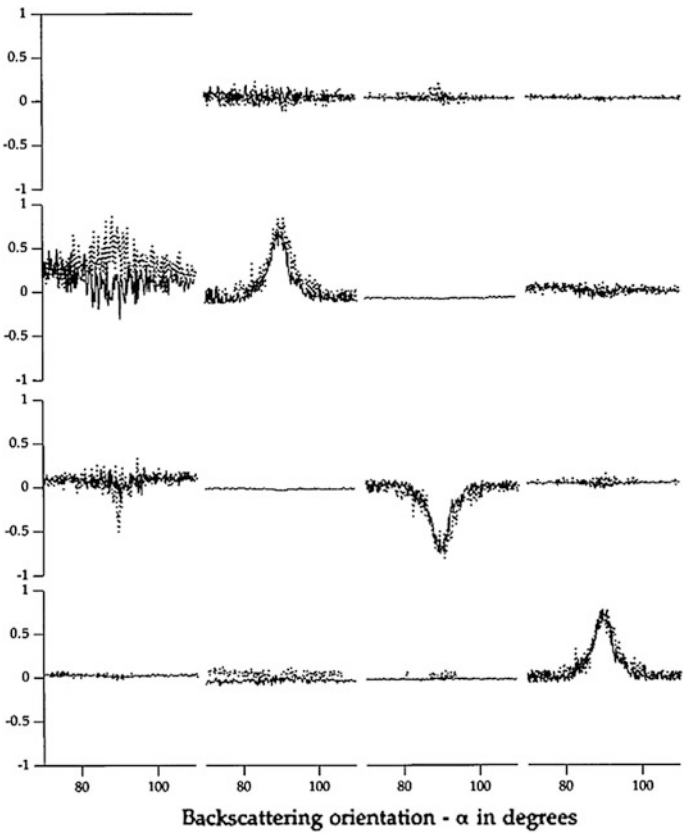


Fig. 10 Normalized Mueller matrix elements as a function of backscattering orientation for L-tartaric acid off- and on-vibration resonance (*solid line* 1029.9 cm^{-1} and *dotted line* 1082.3 cm^{-1}, respectively) (after Carrieri et al. 1998)

Figure 10 displays the example of results of the Mueller matrix element measurements for L-tartaric acid. The dotted curves represent data acquired with a laser line coincident with an absorption band in L-tartaric acid $(1082.3\ cm^{-1})$; the solid curve data were taken at a wavelength off-resonance $(1029.9\ cm^{-1})$. The data in Fig. 10 are qualitatively typical of all the samples. The dependence on the orientation angle is observed only in the diagonal elements and the rapid and seemingly random fluctuations in amplitude as orientation angle is scanned are not electronic detector noise—the pattern is reproduced in detail as the orientation angle is scanned back and forth.

The DIAMMS approach requires that the detection of the suspected contaminant is based on the Mueller matrix elements whose values differ when the probe beam is switched between on- and off-resonance wavelengths at least not lesser than the sum of the standard deviations associated with each of the two measurements. The

Table 4 Measured polarization signatures of L-tartaric acid

Matrix element	Off-resonance 1029.9 cm^{-1}	On-resonance 1082.3 cm^{-1}	Scaled difference Δm_{ij}
m_{13}	0.058	0.188	0.105
m_{21}	0.208	0.704	1.587
m_{31}	0.120	−0.310	−1.101
m_{34}	−0.039	0.131	0.099
m_{43}	−0.038	0.011	0.003

analysis of the data presented in Fig. 10 indicates that, for L-tartaric acid, five Mueller matrix elements correspond to this requirement, see Table 4.

The column "Scaled difference" in Table 4 presents a quantity related to the simple difference between matrix elements on- and off-resonance and is computed by

$$\Delta m_{ij} = \left(\Delta m_{ij}(\alpha, \lambda_o, \lambda_r) - \left\langle \Delta m_{ij}(\alpha, \lambda_o, \lambda_r) \right\rangle\right) / 4 \cdot \mathrm{SD}, \tag{72}$$

where $\Delta m_{ij}(\alpha, \lambda_o, \lambda_r) = m_{ij}(\alpha, \lambda_o) - m_{ij}(\alpha, \lambda_r)$ and the subscripts o and r refer to off- and on-resonance wavelengths, respectively; $\langle\ \rangle$ is the average; SD denotes standard deviations for matrix elements.

Figure 11 is an example of DIAMMS data of biological aerosol protein chicken egg *white albumen* irradiated at beam wavelengths $\lambda = 9.155$ and $\lambda^* = 10.458$ μm.

Normalized and averaged ordinate and abscissa of plots are $m_{ij}^o(\lambda) = \left\langle m_{ij}(\lambda)/m_{11}(\lambda) \right\rangle$ and $m_{ij}^a(\lambda^*) = \left\langle m_{ij}(\lambda^*)/m_{11}(\lambda^*) \right\rangle$, respectively, the intersection of ordinate and abscissa red dotted lines at $\left(m_{ij}^o, m_{ij}^a\right)$ is average M—elements coordinates, and separation widths of paired ordinate and abscissa green solid lines are standard deviations in $\left(m_{ij}^o, m_{ij}^a\right)$ datasets. The perpendicular distance from coordinate $\left(m_{ij}^o, m_{ij}^a\right)$ to the diagonal line in each M—element block is

$$\rho_{ij} = 2^{-1/2}\left(\left(m_{ij}^o\right)^2 + \left(m_{ij}^a\right)^2\right)^{1/2}\left[\cos\left(\arctan\left(m_{ij}^a / m_{ij}^o\right)\right) - \sin\left(\arctan\left(m_{ij}^a / m_{ij}^o\right)\right)\right] \tag{73}$$

where $i, j = 1, 2, 3, 4$ and $\neq 11$. Quantities ρ_{ij} are measures of analyte susceptibility and key identification parameters of the dual-beam-interrogated aerosol. In the example of Fig. 11, for instance, eight most significant values of ρ_{ij} for chicken egg white albumen are grouped in the Mueller space (m_{14}, m_{24}, m_{31}, m_{32}, m_{33}, m_{34}, m_{42}, m_{43}) denoted by solid rectangles.

Some common conclusions relating the results have to be noted (i) a complete Mueller matrix is not required for establishing biosimulant materials features; (ii) diagonal matrix elements undergo the error test rarely; (iii) matrix elements m_{12},

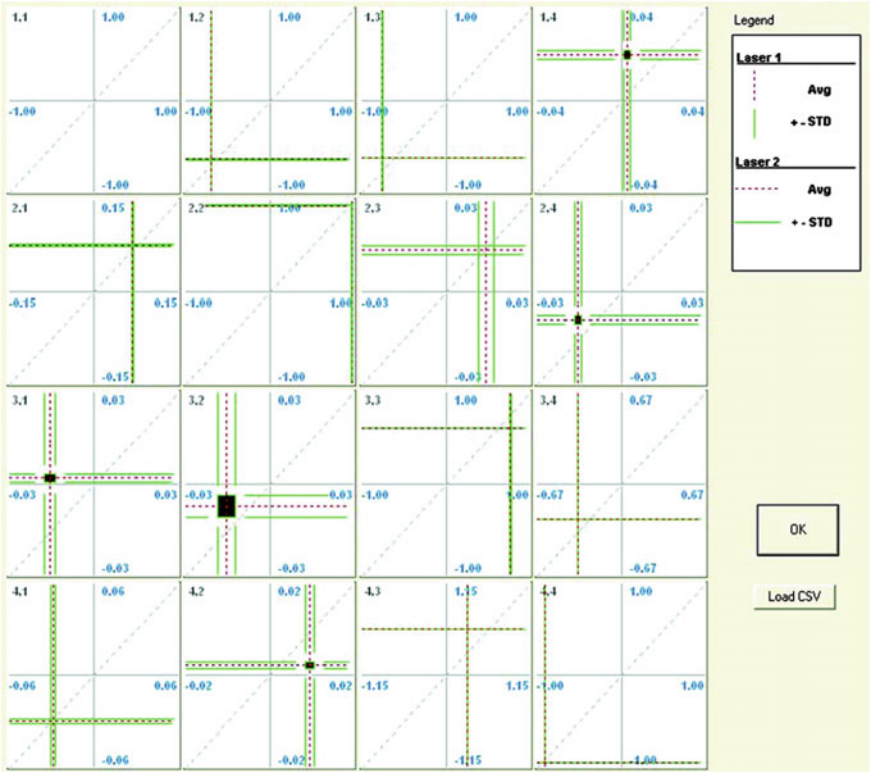


Fig. 11 Mueller matrix elements of biological aerosol protein *chicken egg white albumen*. Mantissa, measurement at $\lambda = 9.155$ μm; abscissa, measurement at $\lambda^* = 10.458$ μm (after Carrieri et al. 2010)

m_{21}, m_{13}, m_{31}, m_{24}, m_{42}, m_{34}, and m_{43} seem to be the most informative detection matrix elements.

5.2 *Urban Objects*

In (Dong et al. 1996) the comparison of two groups of buildings in the test site of Sydney (complete polarimetric data acquired in Sydney, Australia, in August–September 1993 at P band) has been carried out. The motivation was to study the features of polarization signatures of two groups of buildings, which are similar to one another, but are seen for radar at different angles: one group in the near range at 30° and the other in the far range at 60°. Figure 12 shows measured P band polarization signatures (top row) for these two groups of buildings: (a) at incidence angle of 30° and (b) at incidence angle 60°.

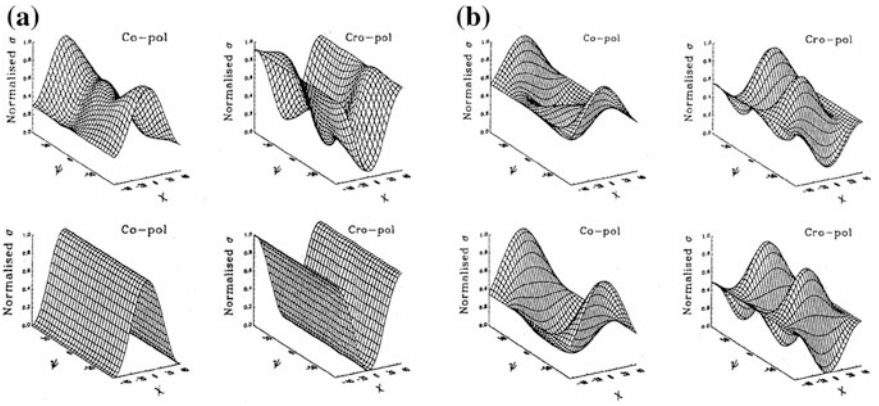


Fig. 12 P band polarization signatures for two groups of buildings: **a** at incidence angle of 30° and **b** at incidence angle 60° (after Dong et al. 1996)

Table 5 Scattering components as a percentage of the total *HH* backscattering response

Incidence angle	Single, %	Double, %	Bragg, %	Error, %
30°	65	35	0	3.4
60°	25	75	0	3.2

For analysis of polarization signatures measured the approach to decompose them into a combination of contributions by four basic scattering mechanisms was suggested. These four basic mechanisms are (see Sect. 4 for details): (i) double bounce scattering; (ii) Bragg scattering; (iii) single (odd) bounce scattering, and (iv) cross-scattering. Bottom row of Fig. 12 presents the theoretic polarization signatures by (a) single bounce and (b) double bounce scattering for comparison.

Table 5 lists the percentages of single, double, and Bragg scattering components of *HH* polarization for two groups of buildings considered at P band. Last column named "error" presents the error between measured and predicted by decomposition suggested values.

The results obtained indicate that at incidence angle 30° single bounce scattering is a dominant mechanism, while, at 60° is a pronounced double bounce scattering. The reason is that the roofs of most of buildings covered by plain tiles are tilted about 30°–35° from vertical. So, in the near range at incidence angle 30° the pronounced specular reflections from the roofs are expected. In far range the roof are not almost normal to the radar and strong double bounce scattering from wall-ground structures is observed.

Identification of man-made objects from radar images has always been a difficult task, especially for single polarization radar. Fully polarimetric radar can provide detailed information on scattering mechanisms that could enable the targets to be identified and classified. However, physical interpretation of scattering signatures remains a challenge resulting from overlaps of single bounce scattering, double

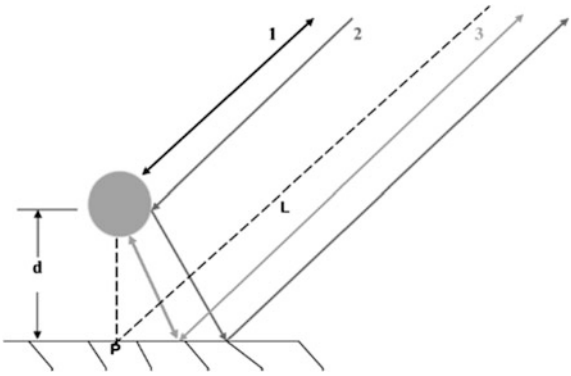


Fig. 13 Paths "1" and "2" are single bounce and double bounce, respectively. Path "3" is triple bounce scattering (after Lee et al. 2006)

bounce scattering, and triple- and higher order bounce scattering from various components of a man-made object, Fig. 13.

In Lee et al. 2006 an interesting example of using PolSAR data to differentiate multiple bounce scatterings is presented. The side under consideration was the Great Belt Bridge, Denmark. Two sets of C band EMISAR data were considered to extract the scattering characteristics of the bridge deck, bridge cables, and supporting structures: the first is obtained during the bridge's construction, Fig. 14a, and the second after its completion, Fig. 14b. The radar look angle is between 28° and 64°, and the range and the azimuth resolution are about 3 m.

Figure 15 shows the results of $|HH|$ for the site under consideration during its construction. Looking at Fig. 15, one would think that the top of the arcs are returns from the two giant cables, the middle two bright lines represent returns from the decks, and the arcs below the two bright lines are double bounce returns from the two cables. However, an aerial photo in Fig. 14a demonstrates that the deck was not installed during that time.

The Pauli decomposition of the PolSAR data corresponding to the case of Fig. 15 is presented in Fig. 16a. The decomposition uses $|HH - VV|$, $|HV|$, and $|HH + VV|$ as red, green, and blue, respectively, and separates the dihedral, cross-pol, and surface scatterings. The absolute value of a polarization orientation angle during bridge construction is shown in Fig. 16b.

It can be seen that the two bright lines in the middle are double bounce from the cables, and the lower arcs are triple bounce returns from the cables. The two giant cables were assembled from several hundreds of small cables, and they were not tightened together by wrapping wires.

The ocean surface has near-zero orientation angle values, whereas, the tilted cables demonstrate higher orientation angle values.

The single bounce from the ocean surface possesses the typical Bragg resonant scattering shown in blue in Fig. 16b. The single bounce from the cables is specular and is shown in green in Fig. 16b. The bridge towers induce a very weak surface backscattering and cannot be discerned.

Fig. 14 Aerial photos of the Great Belt Bridge, Denmark: **a** during the construction and **b** after completion (after Lee et al. 2006)

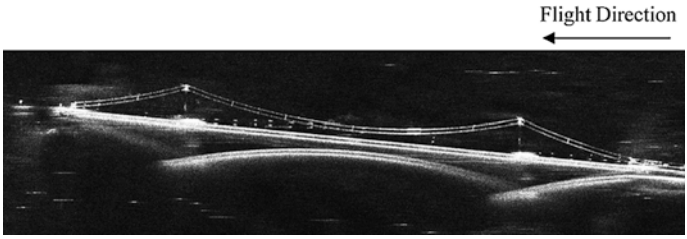


Fig. 15 $|HH|$ images of the Great Belt Bridge, Denmark during the construction (after Lee et al. 2006)

Two middle straight lines in Fig. 16b are from strong double bounce returns. Since the ocean surface is horizontal and flat, the double bounces from the cables are straight lines. The double bounce scattering is higher than those from the single and triple bounces for all three linear polarizations (*HH*, *HV*, and *VV*).

It can be observed from Fig. 16 that the double bounce scattering from the two supporting towers is extremely strong, because double bounce scattering from all parts of the towers is projected down to the ocean surface.

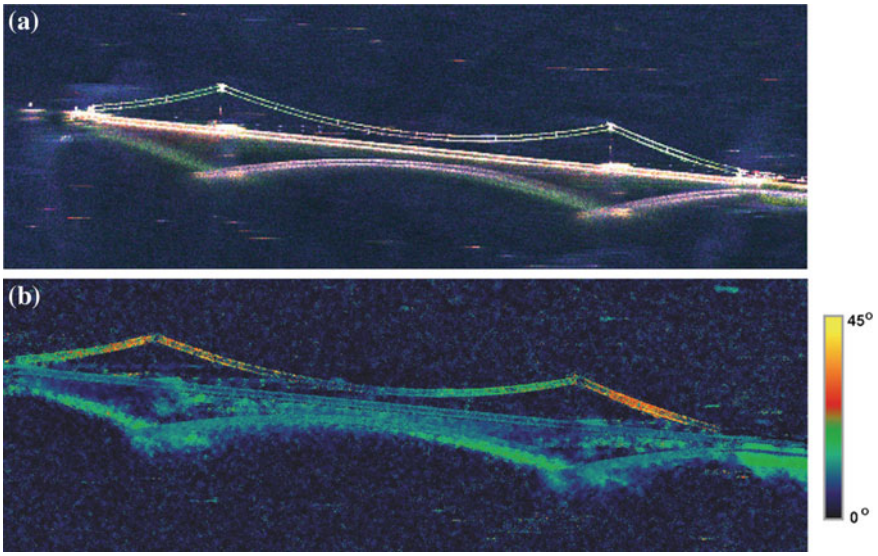


Fig. 16 The Pauli decomposition **a** and polarization orientation angle **b** of the PolSAR data for the Great Belt Bridge during construction (after Lee et al. 2006)

The sections of bridge in green in Fig. 16 have higher values in orientation angles that produce higher *HV* scattering. The other sections are more aligned in the azimuth direction with smaller orientation angles that produce higher scattering in *HH* and *VV*. The scattering from *HH* is higher than *VV*, as shown in Fig. 3, for these sections that produce the near-purple color in Fig. 16. The magnitude of triple bounce scattering is, in general, comparable to that of single bounce scattering.

We see in Sects. 3 and 4 that the Cloude and Pottier decomposition is a good technique to characterize scattering patterns because its derived parameters are rotationally invariant. For example, the entropy is very useful to characterize the diversity of scattering mechanisms, especially for random media.

However, authors report that the entropy is less useful than the alpha angle in characterizing scattering mechanisms for man-made structures like the suspension bridge. The reason is that averaging neighboring pixels (or multilooking) is required to obtain meaningful entropy values, and the amount of averaging and the average of cable pixels with neighboring ocean pixels falsely increases the entropy values around the cables.

While, the average alpha angle is more useful because it can distinguish scattering mechanisms from surface, to dipole, to dihedral, see Fig. 17.

The single bounce scattering from cables shows dipole scattering in green color near 45°. Double bounce from the two supporting towers has an average alpha angle close to 90°. The triple bounce returns are shown in green, implying dipole scattering. We also noticed the absence of local double bounce effect that exists in the triple bounce signature.

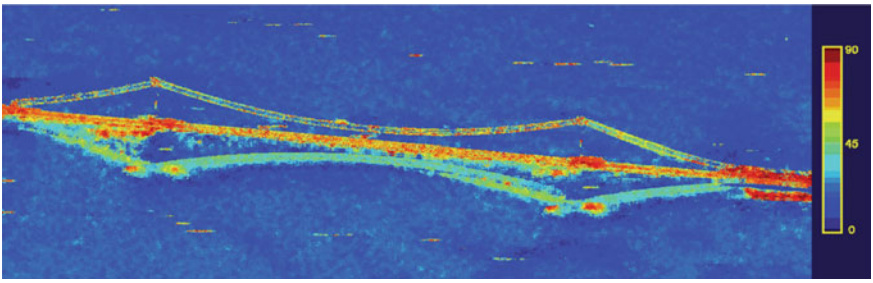


Fig. 17 Averaged alpha angle derived from the Cloude and Pottier decomposition (after Lee et al. 2006)

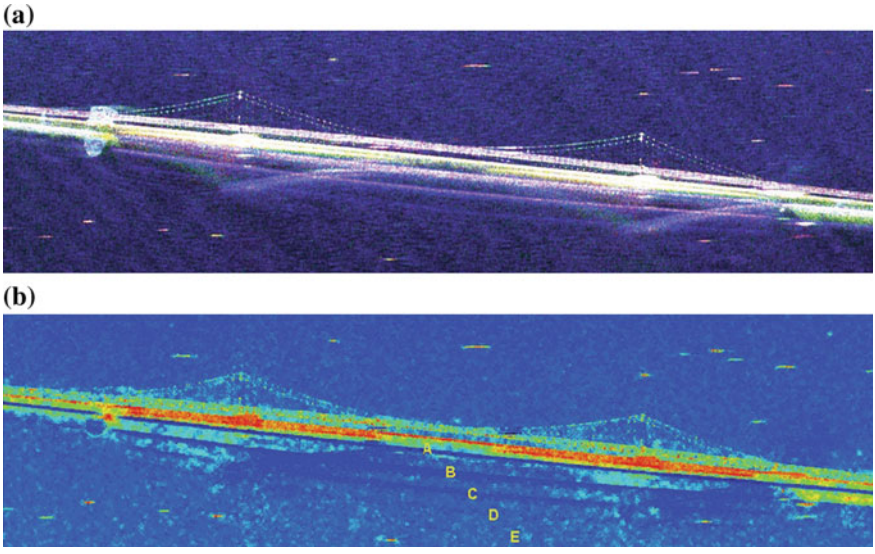


Fig. 18 Pauli decomposition (**a**) and averaged alpha angle (**b**) derived from the Cloude and Pottier decomposition after the completion of bridge construction (after Lee et al. 2006)

The polarimetric signature of the bridge becomes much more complicated after the completion of the construction, see Fig. 17. The deck has been installed, and the cables have been wrapped, reducing their radar cross sections. The addition of the deck makes multiple bounces from the deck overlap with those from the cables.

The Pauli decomposed image, Fig. 18a, shows bridge signatures very differently from those during construction.

The single bounce returns from the deck have a slight curvature as expected, and the scattering is very strong because of the massive structure of the deck. The double bounce from the cables and the deck appear as two straight lines and are totally overlapped as they are all projected down to the ocean surface. The triple

bounce from the cables is weaker and partially obscured by the strong returns from the deck.

The alpha angle image obtained by the Cloude–Pottier decomposition, Fig. 18b, reveals multibounce scattering mechanisms of the cables and the deck much better than that of the Pauli vector Fig. 18a.

Note that the two curvilinear lines, "B" and "C" in Fig. 18b, below the triple bounce signature "A" from the deck also show blue color, which is indicative of an odd bounce in the alpha angle image, but the number of bounces must be higher than three because the slant ranges are longer than that from the triple bounce.

The "C" curvilinear line involves an additional two bounces from the deck to the ocean surface and back to the deck. It can be also observed two more curvilinear lines, "D" and "E" in Fig. 18b, which are somewhat broken up, but still visible in this alpha angle image.

Thus, the results presented have demonstrated the importance of multibounce scatterings, which contribute toward the overall complexity of target signatures.

One more extremely important area where the radar polarimetry can provide the potential capability is the detection of landscape disturbance and damage to man-made objects caused by natural disasters, e.g., earthquake.

In Park et al. (2013) the usability of polarimetric change detection for automatic identification of damage caused by earthquakes and tsunamis is evaluated. Disaster-induced change of backscattering characteristics has been studied using the Phased array L band Synthetic Aperture Radar (PALSAR) onboard the Japanese Advanced Land Observing Satellite (ALOS).

The study site is located in the Ishinomaki area, approximately 90 km west of the epicenter of the 2011 Tohoku earthquake, Fig. 19.

The magnitude 9.0 Tohoku earthquake occurred at 14:46 JST on March 11, 2011, was the worst in Japan's recorded history and one of the five most powerful earthquakes in the world since 1900. It triggered powerful tsunami waves that caused catastrophic damage along the coastal areas of northeastern Japan, with a maximum wave height up to 40 m, Fig. 20a.

A data used to evaluate changes in polarimetric signatures was acquired before 2011 and on April 8, 2011, i.e., about 1 month after the earthquake and tsunami disaster. All datasets used in the study have the same acquisition parameters with the same viewing configuration. Evidently, this allows assuming that differences observed are directly related to the changes that occurred to the ground scatterers.

Second and third rows in Fig. 20b–e present the parameters of Yamaguchi decomposition [surface P_S, double bounce P_D, volume P_V, and helix scattering P_C mechanisms (Yamaguchi et al. 2005)] and of Cloude decomposition (entropy, H, anisotropy, A, and average scattering angle, α, see Sect. 4), respectively.

Evidently, different polarimetric parameters provide different information about earthquake and tsunami damage; therefore, the detectability of each parameter is of great interest. To evaluate detectability of polarimetric parameters, temporal separability between damaged and undamaged conditions is used. Authors suggested several additional parameters for that.

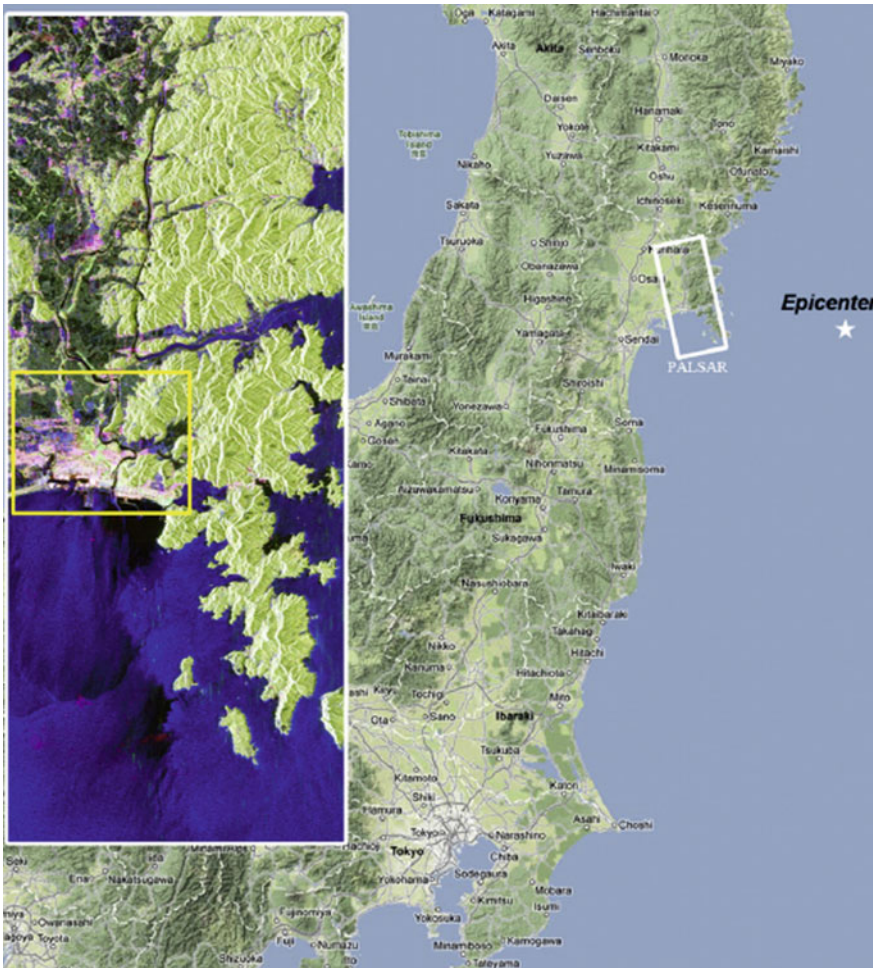


Fig. 19 The Ishinomaki study area and ALOS/PALSAR image acquired after the earthquake (after Park et al. 2013)

The normalized distance, ND, between the mean values of each polarimetric parameter acquired for pre- and postdisaster conditions is defined as

$$\mathrm{ND} = \frac{|m_a - m_b|}{\sqrt{s_a^2 + s_b^2}} \tag{74}$$

where m_a and m_b are the mean values, and s_a and s_b are the standard deviations of pre- and postdisaster conditions, respectively.

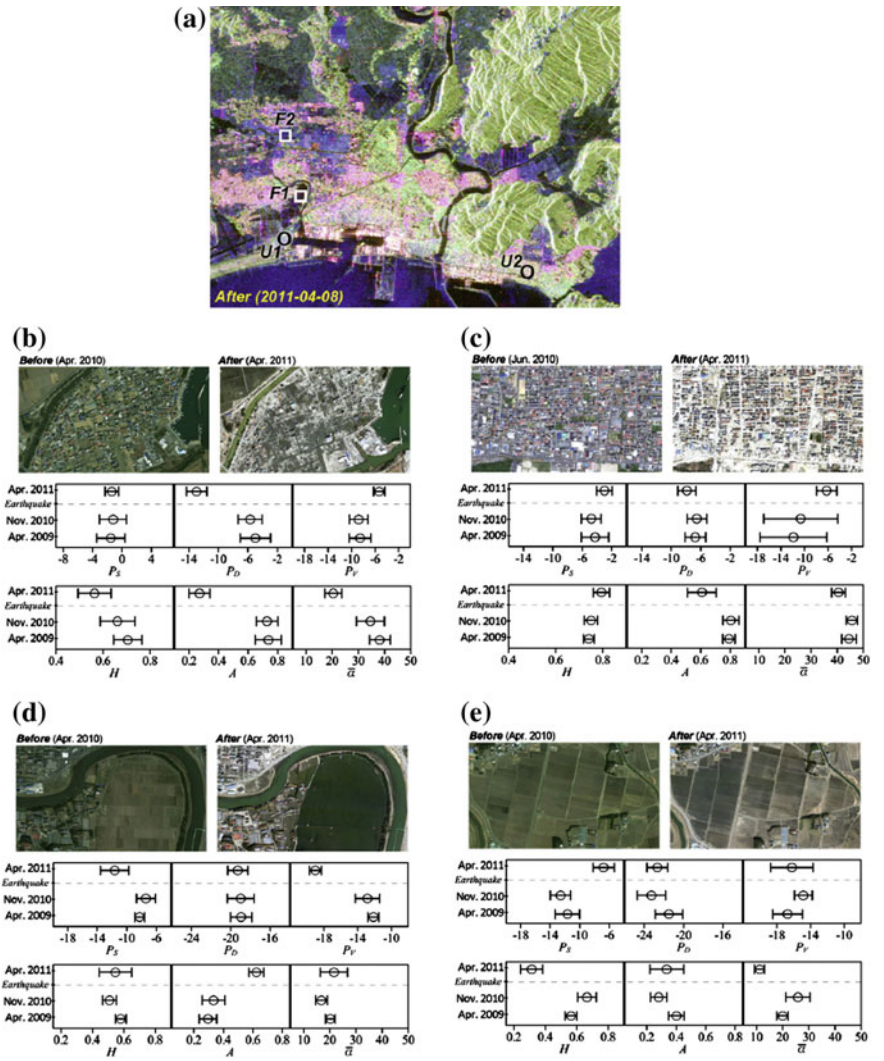


Fig. 20 **a** Image of the study site acquired after (2011/04/08) the earthquake with four subareas of interest (*U1*, *U2*, *F1*, *F2*). Changes in polarimetric parameters in urban areas: **b** totally swept urban area, *U1*; **c** flooded and partly swept urban area, *U2*; **d** heavily devastated area, *F1*, and **e** flooded and partially damaged area, *F2*. *Circle* and *error bar* are sample mean and standard deviation, respectively, of the regions of interest (after Park et al. 2013)

The copolarization coherence in the linear *HV* polarization basis (ρ_{HHVV}) and circular polarization basis (ρ_{RRLL}) and the polarimetric orientation angle (θ_{PO}), are as follows:

$$\rho_{ppqq} = \frac{\left|\left\langle s_{pp} s_{qq}^{*} \right\rangle\right|}{\sqrt{\left\langle s_{pp} s_{pp}^{*} \right\rangle \left\langle s_{qq} s_{qq}^{*} \right\rangle}} \tag{75}$$

$$\theta_{\mathrm{PO}} = \frac{1}{4} \arg\left(\left\langle s_{RR} s_{LL}^{*} \right\rangle\right) \tag{76}$$

Polarimetric coherences have been used to characterize surface roughness (Mattia et al. 1997) and map urban areas (Moriyama et al. 2005), and can be a useful polarimetric discriminator between damaged and undamaged conditions.

The polarimetric orientation angle is the angle about which the antenna has to be rotated in order to align its horizontal polarization axis to the tangential plane of the surface (Schuler et al. 1996). The changes in θ_{PO} may provide information for detecting disaster-stricken building areas, such as inclined man-made objects.

Figure 21 shows change detection maps derived from 2010/11 to 2011/04 pairs for each polarimetric parameter. The reference image in Fig. 6a is derived from the damaged area maps (http://danso.env.nagoya-u.ac.jp/20110311/map/) produced by the Geospatial Information Authority of Japan (GSI).

It is notable that Fig. 21b, c indicates that changes (post–pre) in the conventional single-channel backscatter measurement, for either *HH* or *VV* polarization, cannot provide useful information about damage caused by earthquakes or tsunamis.

As shown in the second row in Fig. 21, ΔP_V exhibits good detectability of damage conditions. In particular, the enhanced class in ΔP_V corresponds to damaged urban areas, while the reduced class can detect part of the flooded farmlands. However, there are false alarms in the undamaged paddy fields in the northwestern quarter of the image due to seasonal changes in land cover type.

ΔH is useful in identifying large areas of flooded farmland. However, significant false alarms can also be found in undamaged areas. ΔA and $\Delta\alpha$ indicate fewer false alarms in seasonally changed areas, although detectable damage is limited to the swept urban area.

$\Delta\rho_{HHVV}$ exhibits the best performance in terms of detectability of both totally and partly damaged urban areas. In this case, however, false alarms can be found in forest areas in the northeastern quarter of the image. $\Delta\rho_{RRLL}$ shows primarily temporal changes in forest areas with limited detectability of swept urban areas. As discussed in the previous section, change detection based on $\Delta\theta_{\mathrm{PO}}$ provides no useful information about damage.

Thus, radar polarimetry has great potential to detect large area of damage. However, it is not always possible to obtain datasets with sufficient spatial and temporal resolutions to assess damaged areas precisely.

The change of radar signal backscattering from swept urban areas is characterized by a significant decrease of the double bounce scattering due to vanished dihedral structures. In partly damaged urban areas, however, such a scattering

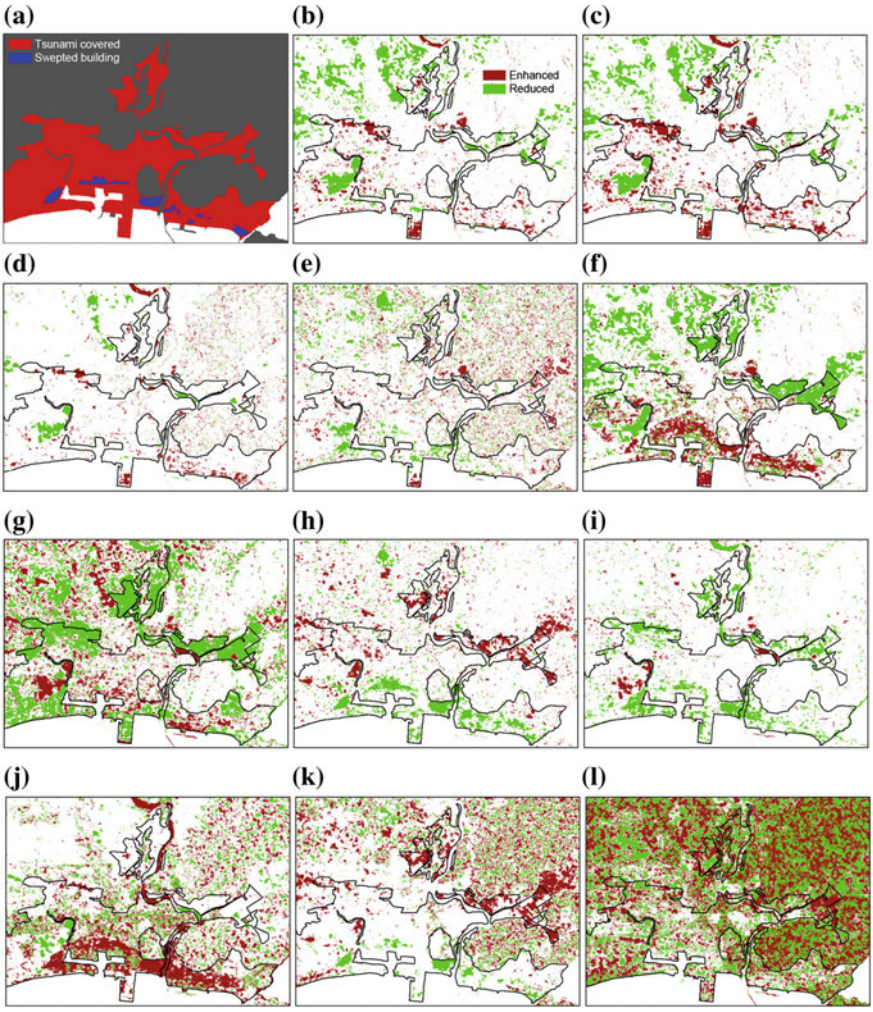


Fig. 21 **a** Reference tsunami damage map, and **b–l** change detection maps derived from the various polarimetric parameters (after Park et al. 2013)

characteristic is not distinctive. Still, minor changes in scattering properties can be identified by polarimetric parameters such as the anisotropy and the polarimetric coherence. On the other hand, changes in scattering properties of farmlands are not characterized easily due to natural or land use dynamics of those areas as well as the delayed post-event acquisition. However, tsunami-induced changes can be detected by the polarimetric entropy or the volume scattering component.

5.3 Mines

Over 50 countries worldwide are contaminated by land mines. Due to high ground and foliage penetrating capabilities, radar polarimetry is a most promising technique for the detection of obscured objects. A key challenge in remote mine detection is discrimination between mines and non-mine objects (e.g., stones). In particular, mines with low metal content are difficult to detect since they can only be found using conventional metal detectors at the expense of high false alarm rates.

To discriminate between mine and non-mine objects, in Hellmann and Cloude (2001) the scattering mechanisms of mines and mine-like objects were analyzed based on the data performed by The European Joint Research Centre (JRC) in Ispra, Italy in a range of frequencies below 5 GHz due to the limited penetration depth for higher frequencies.

On the one hand, since a real mine field may contain anti-personal mines of different types which are not known a priori it is essential to find features which are common for all mines. On the other hand, mines vary in size, shape, and materials, because of that these properties are unlikely to be useful for clutter rejection. However, nearly all types of mines contain an air gap, which will strongly influence the polarimetric signature.

Figure 22 presents the results of α angle and vorticity v, see Sect. 4, for two anti-personal mines TAUPS and MAUS on azimuth (horizontal) and radar frequency (vertical). For oblique incidence, the mines show at discrete frequencies high alpha and vorticity values (called α-resonances). For close to vertical incidence angles the features disappear as expected from symmetry arguments. The features are largely independent of the azimuth angle.

JRC database provides polarimetric measurement data over a wide frequency range for a variety of non-mine targets which are likely to be found amongst mines in a real mine field (e.g., aluminum cans, plastic bottles, stones, wood, etc.). The results of α angle and vorticity v for stone (upper row) and plastic bottle (lower row) on azimuth and frequency are presented in Fig. 23.

Figure 22 shows that the distribution of the alpha and vorticity is a function of the azimuth angle. However, for any specific azimuth the distribution always differs from those of the mines. While the mines show clear resonances, the non-mine targets do not. Only for nonmetallic objects with an air gap inside [e.g., the plastic bottle shown in Fig. 23 (bottom)] do the features show some similarities, since the structure of these objects (nonmetallic cylinder with air gap inside) is similar to the structure of mines.

To evaluate the polarimetric properties of background surface clutters, authors used a set of polarimetric surface measurements provided in Nesti (1998). The results are shown in Fig. 24.

The distributions observed are clearly different from those of the mines and the non-mine targets.

The results shown above indicate that radar polarimetry can provide substantial information for the detection of anti-personal mines as well as for the discrimination

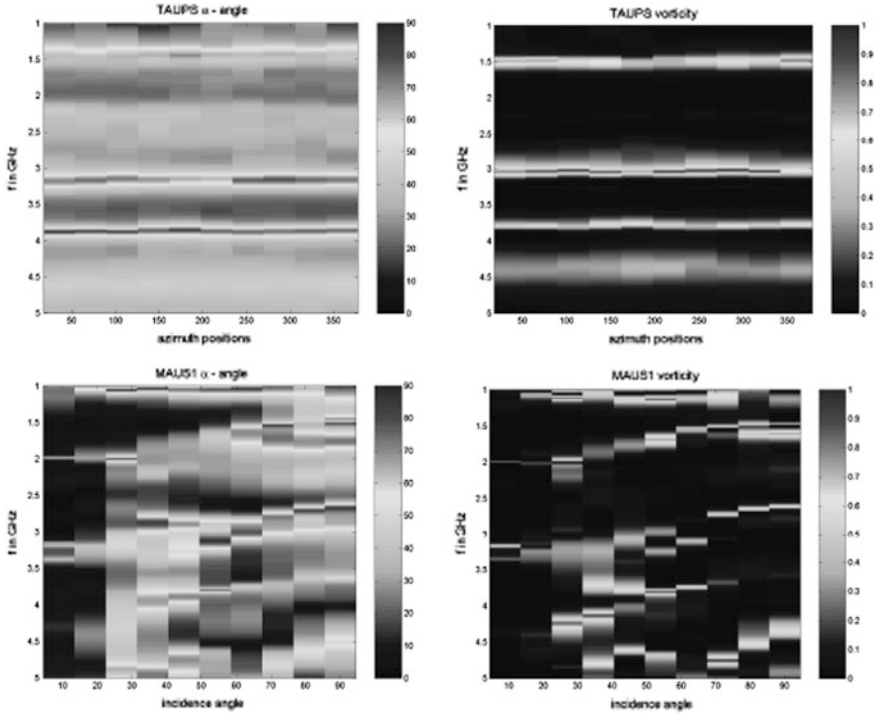


Fig. 22 Dependences of α angle (*left*) and vorticity (*right*) for two mines on azimuth and radar frequency (after Hellmann and Cloude 2001)

between mines, non-mine targets, and background surface clutters at oblique incidence angles. However, in order to validate these findings and also for buried mines more measurements on surface laid mines are necessary. Because of the scattering behavior of the various non-mine targets differ recognizably, further polarimetric measurements on such objects could enhance the understanding of the scattering mechanisms and yield the possibility to discriminate between different types of non-mine targets. Additionally, this information might be useful for larger sensors like airships for the detection of mine fields.

The problem of discrimination between mines and surface clutters is of great interest in optical range as well. In Stabo-Eeg et al. 2008a the reflection spectrum and Mueller matrices of seven different land mines (TMRP-6, TMA-1, TMA-4, PMA-1, PMA-2, PMA-3, and MRUD) and three different plants (*Yucca guatemalensis*, *Plantago major,* and *Taraxacum officinale*) were measured. The reflection spectrum is taken within an integrating sphere (Labsphere DRA-CA-50D) between 400 and 2500 nm, see Fig. 25.

The reflection coefficients are between 5 and 15 % and peak values are around 510 nm for all the mines, which contained green pigment, while PMA-3 made out

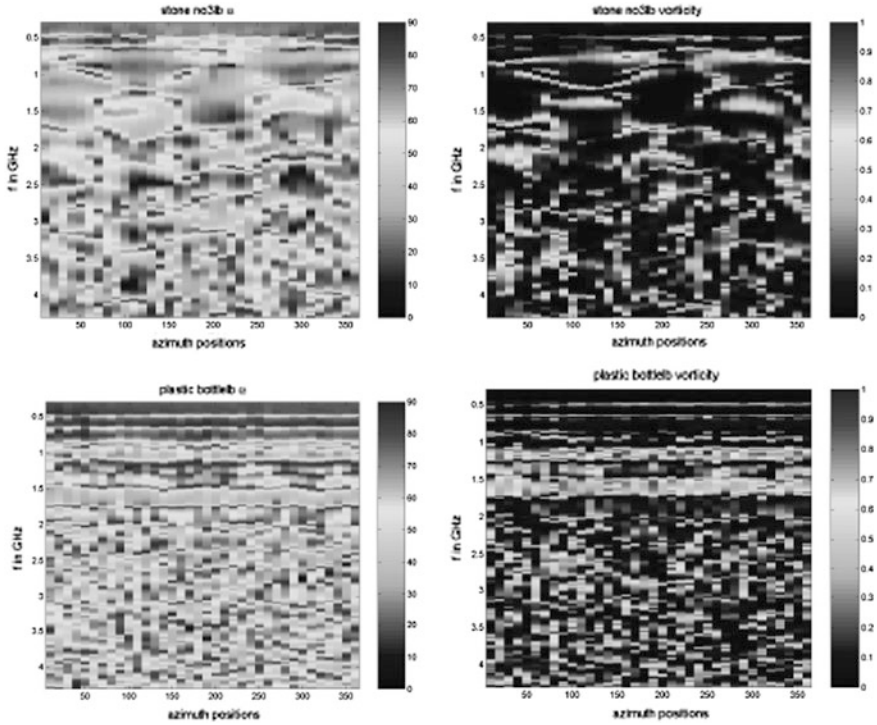


Fig. 23 Dependences of α angle (*left*) and vorticity (*right*) on azimuth and frequency for a stone (*top*) and a plastic bottle (*bottom*) (after Hellmann and Cloude 2001)

of black rubber material is characterized by constant reflection coefficient around 5–6 %.

The mines were found to be less reflective in wet compared to dry conditions, see Fig. 26.

In the polarimetric study, the Mueller matrices from specular and nonspecular reflections were collected from the sample pieces from the cover of seven plastic mines (the pieces were cut out from the covers and had a size of about 3×3 cm) and compared against three different plants (*Y. Guatemalensis*, *P. major,* and *T. officinale*). More detail information on polarimeter used in this experiment can be found in Stabo-Eeg et al. (2008b, c).

As an example, the results of nonspecular Mueller matrix measurements for the mine TMRP-6 and leaf from the plant *Y. guatemalensis* are presented in Figs. 27 and 28, respectively. Measurements were performed at 1570 nm.

The mines clearly show a specular reflection behavior whereas reflection from the plants is dominated by scattering. To compare the depolarization properties of the mines and plants under consideration the depolarization index (Gil and Bernabeu 1986) was calculated from the measured Mueller matrices. The results are presented in Fig. 29

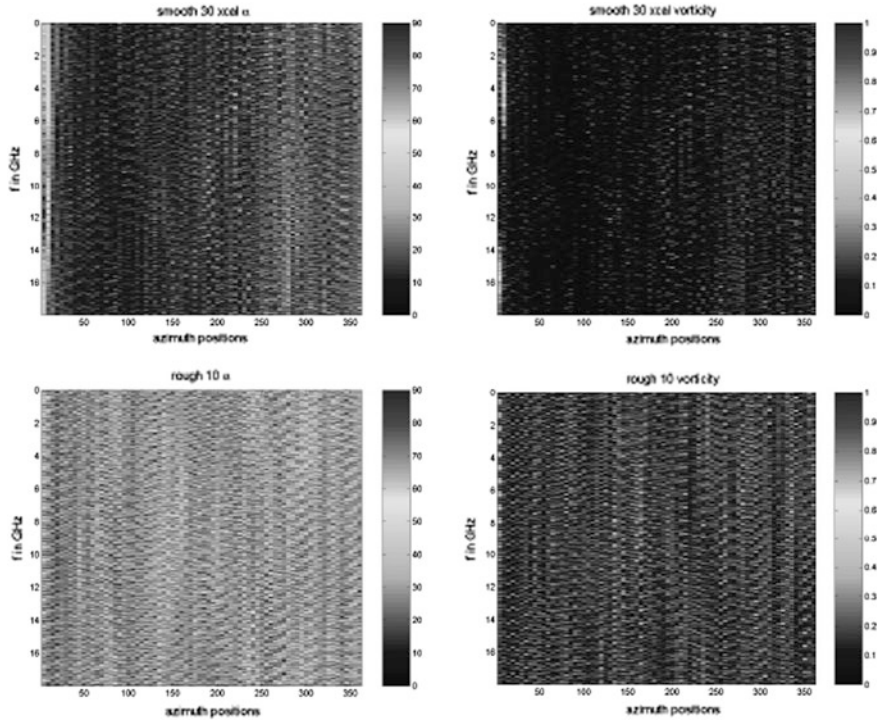


Fig. 24 Dependences of α angle (*left*) and vorticity (*right*) on azimuth and frequency for smooth (*top*) and rough (*bottom*) surfaces (after Hellmann and Cloude 2001)

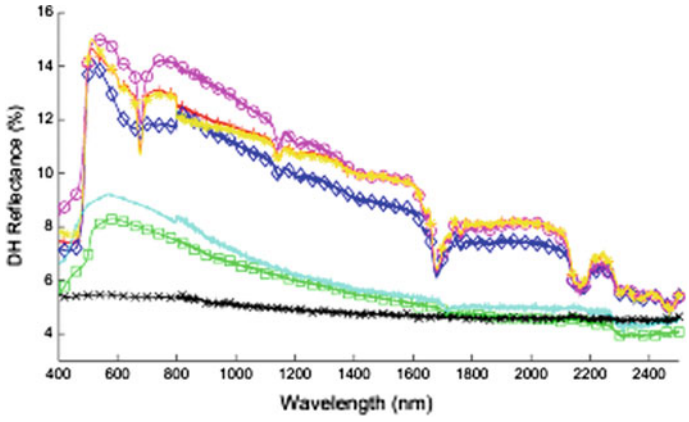


Fig. 25 Spectroscopic reflectance measurements of the mines: (*Black circle*) PMA-1, (*white circle*) PMA-2, (*times*) PMA-3, (*plus*) TMRP-6, (*asterisk*) MRUD, (*square*) TMA-1, (*diamond*) TMA-4 (after Stabo-Eeg et al. 2008a)

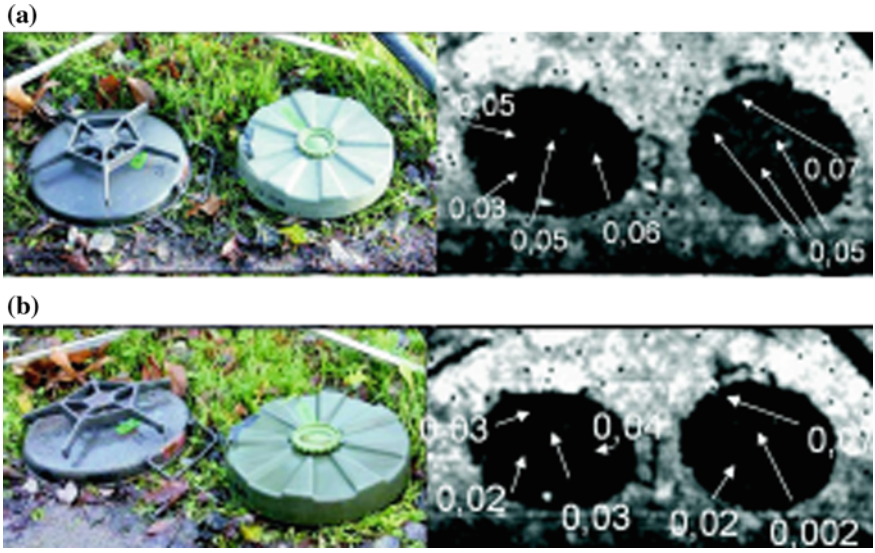


Fig. 26 Digital camera images and scanning laser radar reflectivity of the mines m/47 (*left*) and a TMA-1 (*right*) in: **a** dry condition; and **b** wet condition (after Stabo-Eeg et al. 2008a)

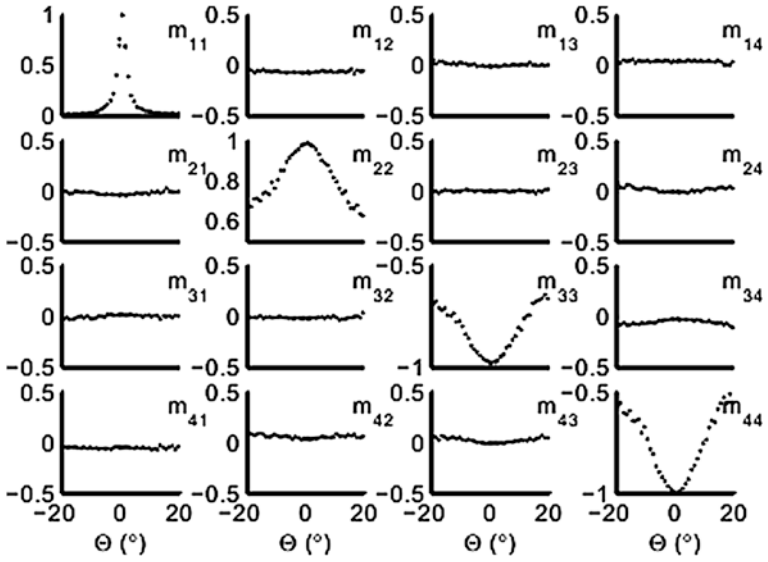


Fig. 27 Dependences of the Mueller matrix elements of the mine TMRP-6 on scattering angle θ. All elements are normalized by m_{11} except m_{11} normalized by the maximal intensity value (after Stabo-Eeg et al. 2008a)

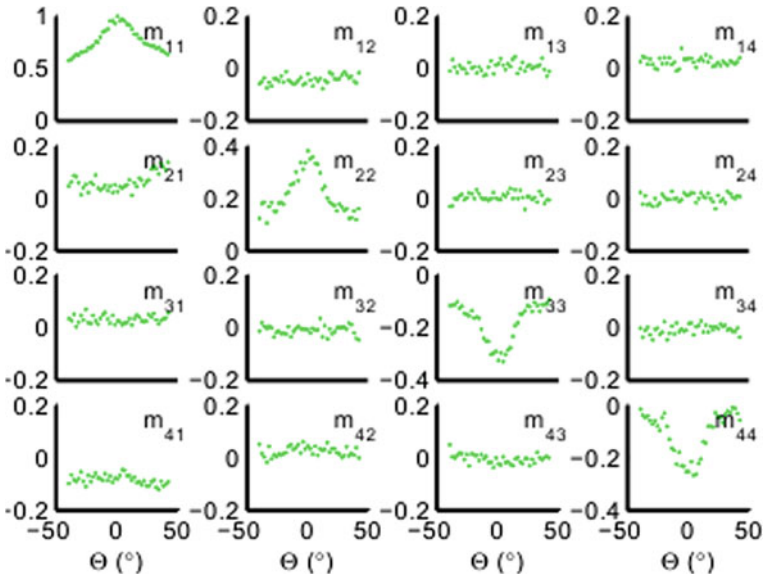


Fig. 28 Dependences of the Mueller matrix elements of a leaf from the plant *Yucca guatemalensis* on scattering angle θ. m_{11} is in arbitrary units (after Stabo-Eeg et al. 2008a)

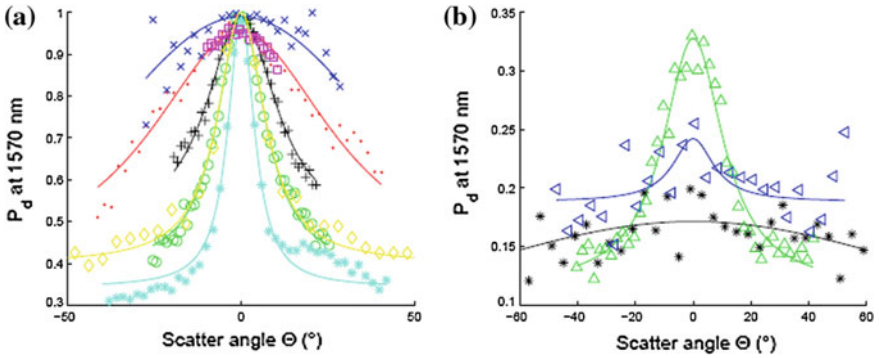


Fig. 29 Dependences of the depolarization index on scattering angle for different samples of mines and plants at 1570 nm. (a) Mines: (*Black circle*) PMA-1, (*white circle*) PMA-2, (*times*) PMA-3, (*plus*) TMRP-6, (*asterisk*) MRUD, (*white square*) TMA-1, (*diamond*) TMA-4; (b) Plants: (*triangle*) *Yucca guatemalensis*, (*left triangle*) *Taraxacum officinale*, (*asterisk*) Plantago major (after Stabo-Eeg et al. 2008a)

As it can be seen, the plants exhibit a larger depolarization than that of the mines. This is an expected result since the plants are partially transparent at NIR wavelength while the mines are more reflective. The depolarization of the plants is probably caused from both scattering inside the leaves as well as different surface

profiles. Other depolarization effects originating from retardation and diattenuation seem to be very small, since the nondiagonal matrix elements are close to zero.

Thus, Mueller matrix measurements and, in particular, depolarization information at NIR seem as well a promising supplementary technique to distinguish surface mines from natural vegetation backgrounds.

5.4 Archaeology

While remote observation for the analysis of archaeological structures has quite long history starting at the end of the nineteenth century when in 1879 the ancient city of Persepolis was studied using aerial photographic documentation, the SAR polarimetric technique is rather new tool of investigation for cultural heritage applications.

Below in this subsection, we examine several examples to assess the potential of radar polarimetry to identify buried archaeological structures. These are sets of L band PolSAR data of three archaeological sites: Pelusium on the northeastern edge of the Nile Delta, Egypt (Stewart et al. 2013); Djebel Barkal the northern province of Meroe, Sudan, and Samarra, 130 km north of Bagdad, Iraq (Dore et al. 2013; Patruno et al. 2013).

The choosing of L band for this analysis is explained by deeper penetration of the SAR signal in dry sand cover overlying potential buried structures. Indeed, if λ is the wavelength and $\varepsilon = \varepsilon' - \varepsilon'$ is the relative complex dielectric constant, for materials with $\varepsilon'/\varepsilon' < 0.1$, then an approximate estimation of the penetration depth of a SAR signal through a volume L_P can be obtained as follows (Ulaby et al. 1982):

$$L_P \cong \frac{\lambda\sqrt{\varepsilon'}}{2\pi\varepsilon'} \tag{77}$$

Equation (77) shows that the greater the wavelength, the greater the penetration depth. Thus, ALOS PALSAR operates in L band (23.6 cm), which is noticeably longer than other current, high-resolution, spaceborne SAR systems, which operate either in C band (5.6 cm) or X band (3.1 cm).

Pelusium (modern Tell el-Farama) is a relatively large area with flat terrain locating at latitude $31^\circ\ 2'23.00''N$ and longitude $32^\circ 33'30.74''E$ on the northeastern edge of the Nile Delta in Egypt at the mouth of the now extinct Pelusiac branch of the Nile. It lies along the "Ways of Horus" and during Roman times it was the most important harbor in Egypt after Alexandria. This area is characterized by a long human occupation and thereby is considered the important archaeological site in northern Sinai.

Samarra was the second capital of the Abbasid Caliphate. It was the capital of Iraq only for 56 years (AD 836–892), when the city was suddenly and mysteriously abandoned. This allowed very good preservation of its topography and provides an important example for the study and comprehension of the ancient urban system for

archaeologists. The portion of the whole area of Samarra properties (15,058 ha) ($N43^{0}45'-43°51'$; $E34°25'-34°05'$) presented in this work is the octagonal city of al-Qadisiyya, an unfinished city built in mudbrick and still unexcavated. It was built, according to the texts, by Caliph Harun al-Rashid on the model of the Round City of Baghdad and abandoned in AD 796, before Samarra city was built.

Djebel Barkal is one of the five Napatan (900–270 BC) and Meroitic (270 BC to AD 350) archaeological sites stretching over more than 60 km in the Nile valley, in a semidesert area considered to be part of Nubia. It was the capital of the Kushite kingdom, probably by the end of ninth century BC, keeping both its religious and administrative roles until the fourth century. This area was considered to be sacred, with its tombs, temples, and pyramids, since New Kingdom times (ca. 1500 BC). Several temples at the foot of the sacred hill and facing the Nile, have been revealed by excavations and survey, together with administrative palaces and pyramids.

The 18 PALSAR images acquired in various polarimetric modes were obtained for the analysis through a European Space Agency Category-1 project. Processing included coregistration and summation of images for each polarimetric mode in order to reduce speckle, see Figs. 30 and 31. The summed images were then

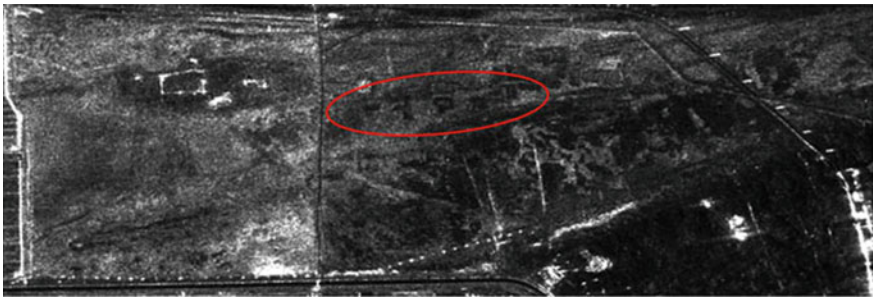


Fig. 30 Sum of eight PALSAR images acquired from 2007 to 2010 in single polarization mode (*HH*). The red ellipse indicates features of potential archaeological interest (after Stewart et al. 2013)

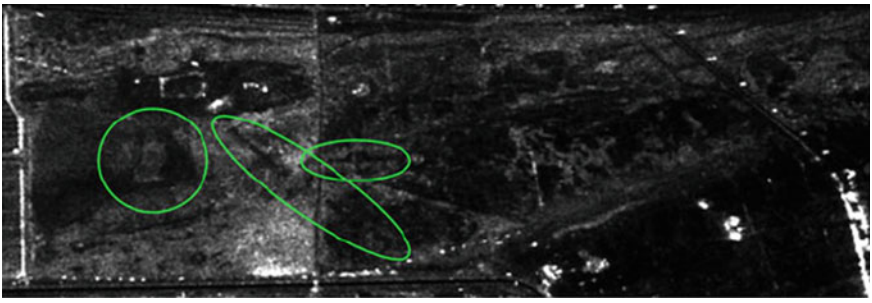


Fig. 31 Sum of nine PALSAR images acquired from 2007 to 2010 in dual polarization mode (*HH*). Green ellipses indicate features of potential archaeological interest (after Stewart et al. 2013)

Fig. 32 Optical image acquired on August 1, 2009 (after Stewart et al. 2013)

compared with each other and with optical imagery, Fig. 32 (Courtesy Google Earth. Image copyright 2013 GeoEye. Copyright 2013 ORION-ME), to identify any features that may be of archaeological interest.

Comparing the *HH* polarized images, some features are more evident in one than the other, the reasons probably being simply due to the different acquisition dates of the images in each stack, and the differences in spatial resolution and pixel spacing. Comparing the SAR images with the optical, Fig. 32, many of the features in the SAR are visible also in the optical imagery, but in some cases are not so clearly delineated.

A number of what appear to be paleaoenvironmental features are highlighted by the SAR dual *HH* polarization images shown in Fig. 33, where arrows indicate both brighter and darker radar features.

Yellow arrows indicate possible Nile palaeobranches, whereas blue and red arrows indicate potential palaeochannels. The potential palaeochannels highlighted by the blue arrows are spatially linked to the area of urban remains. Previous studies have suggested the presence of minor northern branches of the Nile, easily identified in Fig. 33 by red arrows, the silting up of which occurred around AD 25, as suggested by Sneh and Weissbrod (1973). Ancient ruins have been found very close to these possible palaeoriver courses, and therefore, the potential channels fit well with these findings. The red arrows in Fig. 33 indicate both brighter and darker features and, in particular, darker features that connect the defunct river branches with the baths, thus suggesting that it may be related to a Roman hydraulic system built to bring water to the thermal areas.

Results of the polarimetric analysis deriving Freeman decomposition [odd bounce (red), volume scattering (green) and double bounce (blue)] and entropy undertaken on the SAR quad-pol image of Pelusium archaeological site are shown in Fig. 34.

Figure 34a shows that the dominant scattering mechanism over the site is single bounce. This means that the incident SAR signal generally scatters once on the surface before returning to the sensor and does not generally get depolarized. Bright pixels have high entropy, indicating mixed scattering. Dark pixels have low entropy, indicating the presence of a dominant scattering mechanism, Fig. 34b.

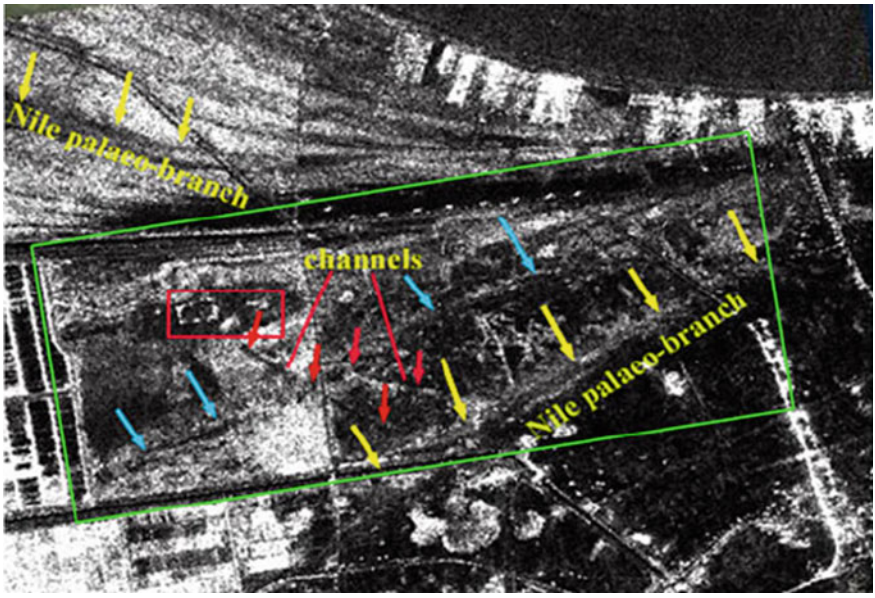


Fig. 33 Sum of two SAR dual *HH* polarization images acquired on Jul 28,2008 and Sep 15, 2009 (after Stewart et al. 2013)

Green and blue ellipses circle highlight features of potential archaeological interest (Image provided by the European Space Agency.).

In Fig. 35 an example of entropy and alpha angle from the al-Qadisiyya area together with H/α plane is shown.

High values of H and medium values of α are recorded in almost the whole of the zone corresponding to surface roughness (Zone 8) and vegetation (Zone 5) in the H/α plane. Except for part of the octagonal city walls and part of interior system of the canals (black ellipses in the image), both recorded with low values of H and low values of α, no other archaeological features were detected by the ALOS PALSAR sensor (Dore et al. 2010).

The low archaeological information gained with H and α polarimetric descriptors turned our attention to other parameters in order to test the validity of the scattering mechanisms both upon the already-known structures and in the areas around them.

The results of Freeman (a) and Yamaguchi 3 (b) decompositions in the form of RGB images (R—double bounce; G—volume scattering; B—single bounce) are presented in Fig. 36.

From an archaeological point of view, the features, which can be detected in Fig. 36, are also well known in optical images and previously recorded in the literature. The wall perimeter is well detected and visible in the Freeman image (white arrows). It responds to a single bounce scattering mechanism. Note that there are no responses from the qanāt, the underground channel entering the city through the northwest wall with clear surface alterations. Authors assumed that even if at a

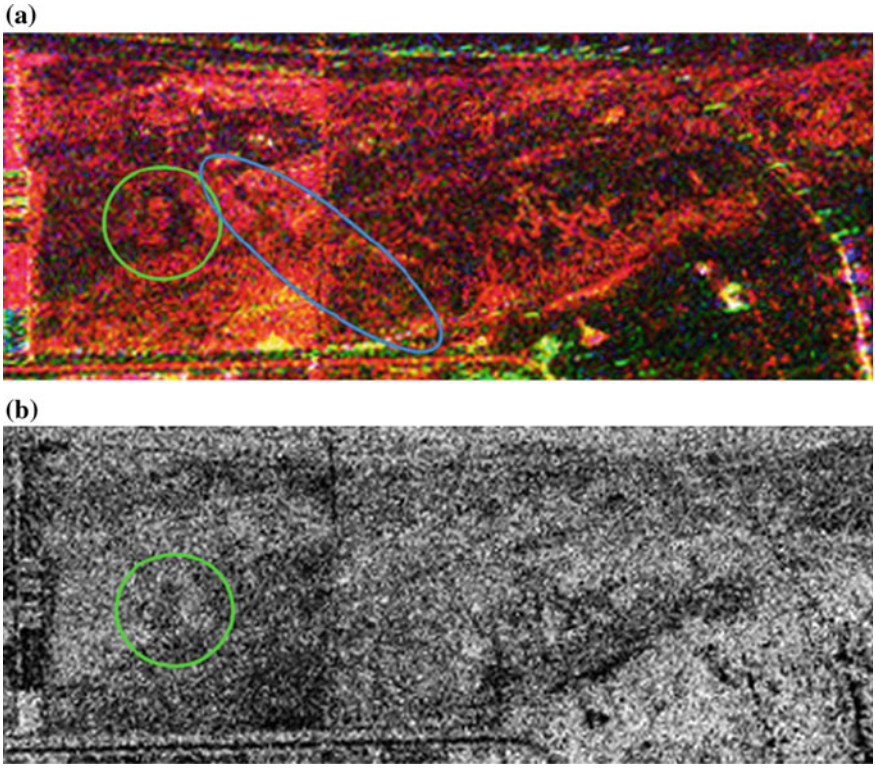


Fig. 34 **a** Freeman decomposition and **b** entropy for SAR quad-pol image acquired on April 8, 2009 (after Stewart et al. 2013)

first level of interpretation this absence of response may be ascribed to the low ALOS PALSAR spatial resolution (about 20 m), however, it is hardly possible as the width of the surface alterations is about 64 m wide. Most of all, it is possible to detect the qanāt in ENVISAT-ASAR imagery (a series of images spanning from 2004 to 2008) with a spatial resolution of about 30 m. These results therefore are not due to their spatial resolution, but to the deformation of the surface.

Concerning Djebel Barkal, the same polarimetric decompositions were performed. The area investigated is the portion northwest of the archaeological complex, near the first group of Royal pyramids. Here, a strong backscatter is noticed in all the products obtained: Fig. 37a—Pauli RGB (Patruno et al. 2013); Fig. 37b—Freeman; and Fig. 37c—Yamaguchi decompositions.

The additional analysis including the 3 years (2006 and 2009 years) difference in time acquisition shows that only for that particular bright response there is a correspondence in all images. It is interesting that the target is visible in an area where a high content of archaeological structures is present, but it is not visible in optical

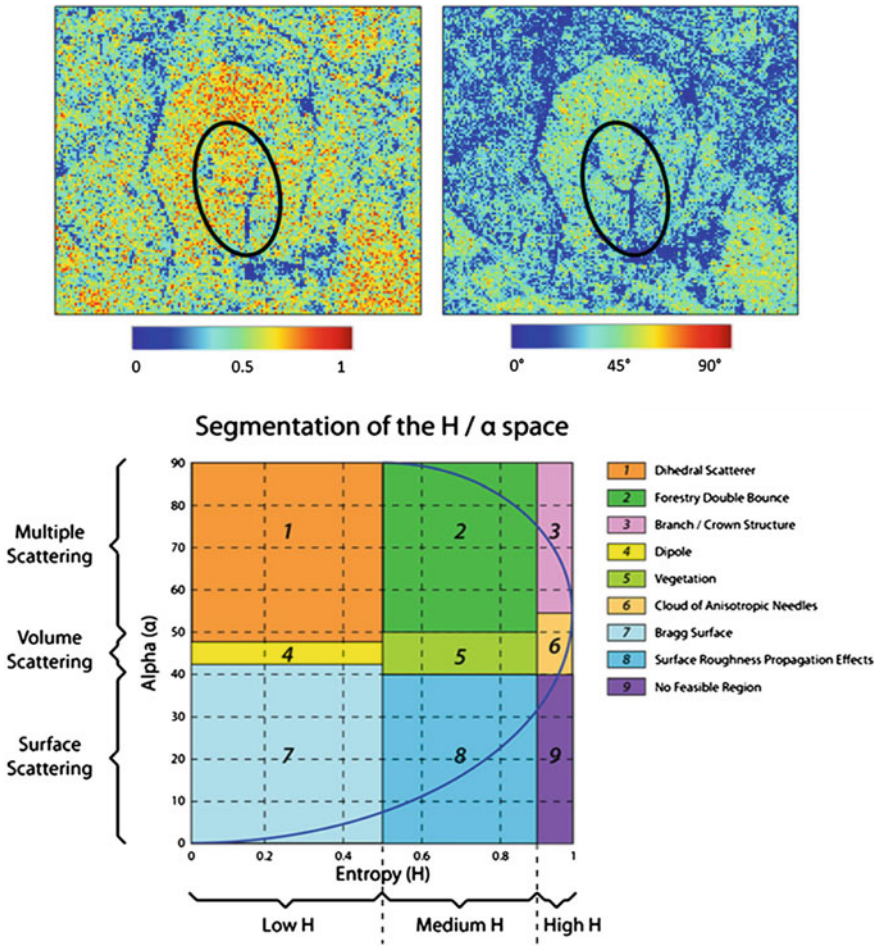


Fig. 35 Entropy (*left*) and alpha angle (*right*) and the H/α plane (*below*) (after Dore et al. 2013)

images, see Fig. 38. This could be due to penetration of radar waves. Latter it is confirmed also by Yamaguchi decomposition, Fig. 37c.

In the ellipse, Fig. 38, is the area where the anomaly was detected in Fig. 37. In rectangles, the three archaeological areas where remains are already known are marked.

Thus, the analysis conducted above with the SAR data of archaeological areas under consideration results in the identification of a number of what appear to be palaeolandscape and archaeological features. However, validation of these findings, possibly by geophysical survey or excavation, would be required to determine the precise nature of these anomalies. On the other part, SAR data analysis demonstrates great potential for places of cultural interest where a validation or archaeological excavations in situ are not allowed for a variety reasons or have not yet

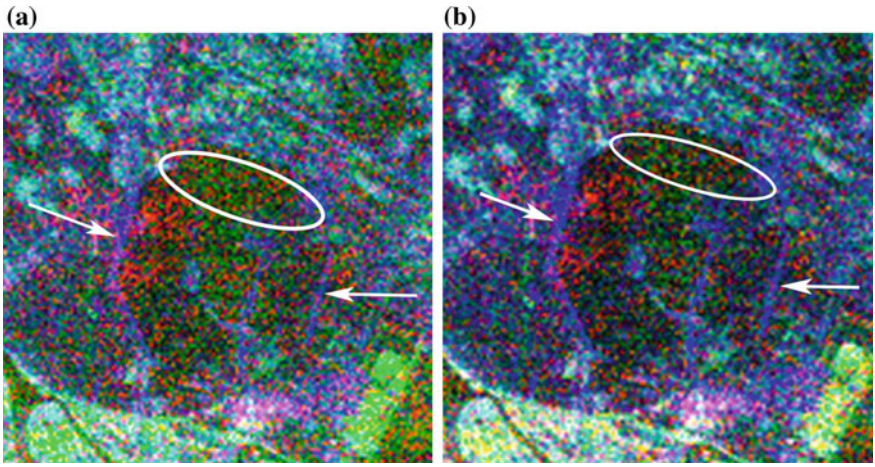


Fig. 36 Freeman (**a**) and Yamaguchi 3 (**b**) decompositions for the Samarra archaeological area (al-Qadisiyya city) (after Dore et al. 2013)

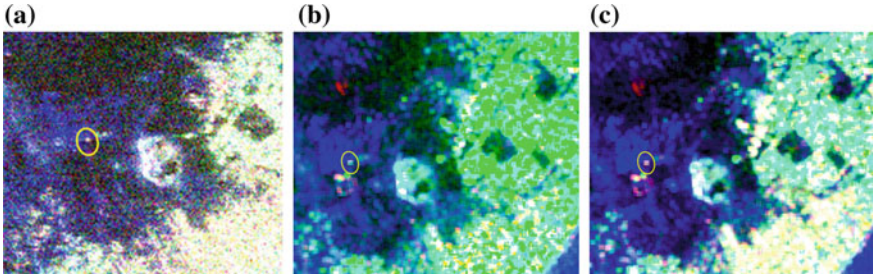


Fig. 37 **a** Pauli decomposition (after Patruno et al. 2013), **b** Freeman and **c** Yamaguchi 3 decompositions (after Dore et al. 2013). The *yellow ellipse* outlines the persistence of the strong scattering noticed in the area

started. Moreover, this allows monitoring the threats and risks caused by uncontrolled urbanization, agricultural exploitation, or even illegal excavations.

The spatial resolution of the SAR data does not permit the identification of small-scale features. However, the features can be clearly distinguished in SAR data at a scale at which the same features may be very difficult to distinguish in optical imagery. This makes radar polarimetry a useful tool to be applied in a preliminary observation of ancient areas providing a priori information about the natural and human environment and an overview of the historical zones.

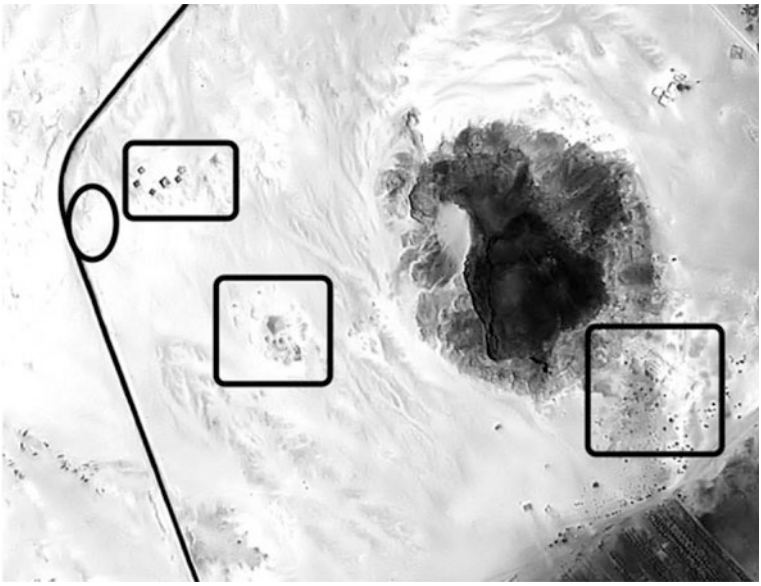


Fig. 38 Optical image (Kompsat-2) image of Djebel Barkal archaeological area (after Patruno et al. 2013)

5.5 *Miscellaneous Man-Made Targets*

Remaining in the scope of the primary motivation for research in scatter polarimetry, i.e., to gain understanding of the physical features of the interaction of polarized beams with natural scenes and to search for useful discriminants to classify objects at a distance, the main subject of interest in this section is polarimetry of various man-made objects as diverse as wood samples and automobile paints.

In Moriyama et al. 1997 radar polarimetry was used to examine classification of target, metallic plate, and pipe, buried in the underground. Measurements of the scattering matrix, Fig. 39, are carried out at the Niigata University Campus in the conventional linearly polarized HV basis and from 250 MHz to 1.0 GHz. The targets are a metallic plate of 25 × 25 cm dimension and a metallic pipe of 10 cm in diameter and 100 cm long which was buried in the ground at the depths of 50 and 80 cm, respectively, see Fig. 39.

Further, for classification of buried target authors used some polarimetric identifiers extractable from the measured scattering matrices: the polarization anisotropy coefficient, polarimetric signature, etc.

The results for polarimetric power signature of targets under consideration are displayed in Fig. 40.

Note, that the tilt angle of co-pol maximum polarization state corresponded to the orientation angle of buried cylindrical target. The polarization ratio of maximum in Fig. 40b is found to be $\chi_{\max} = 1.158 - i0.236$, which corresponds to the

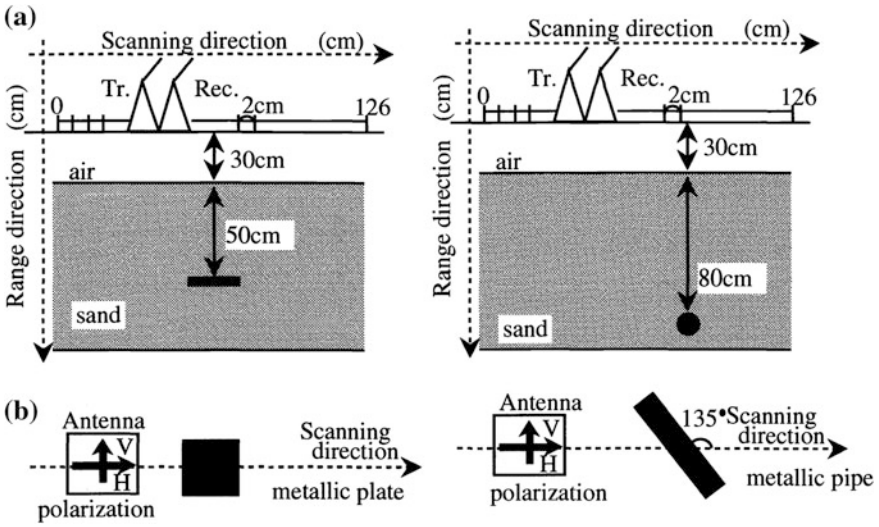


Fig. 39 Schemes of experiments (after Moriyama et al. 1997)

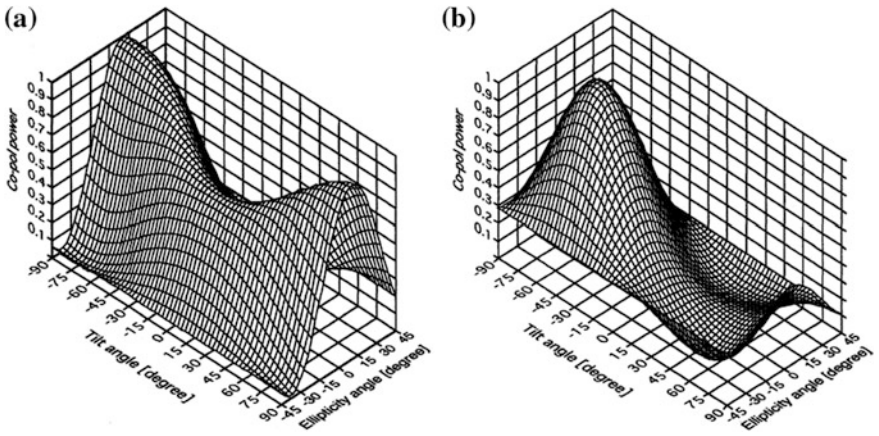


Fig. 40 Polarimetric copolar signatures of **a** plate and **b** pipe (after Moriyama et al. 1997)

characteristic polarization states of $\theta = -49.9°,\ \varepsilon = -5.7°$ for the pipe. So, the tilt angle of the co-pol maximum polarization state is quite close to the actual target orientation.

The power polarization anisotropy coefficient is defined as

$$\mu = \sqrt{1 - 4\left|\frac{\mathrm{Det}(\mathbf{S})}{\mathrm{Span}(\mathbf{S})}\right|^2} = \frac{|\lambda_1|^2 - |\lambda_2|^2}{|\lambda_1|^2 + |\lambda_2|^2}, \quad |\lambda_1| \geq |\lambda_2|, \quad 0 \leq \mu \leq 1, \tag{78}$$

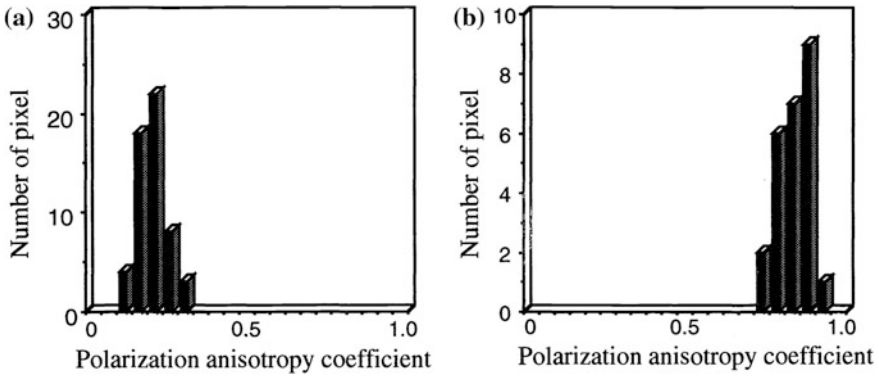


Fig. 41 Distribution of power polarization anisotropy coefficients of **a** plate and **b** pipe (after Moriyama et al. 1997)

where λ_1 and λ_2 are the eigenvalues and $\mathrm{Det}(\mathbf{S})$ and $\mathrm{Span}(\mathbf{S})$ are determinant and span of the scattering matrix Eq. (55)

$$|\mathrm{Det}(\mathbf{S})| = \left|s_{HH}s_{VV} - s_{HV}^2\right| = |\lambda_1\lambda_2| \tag{79}$$

$$\mathrm{Span}(\mathbf{S}) = |s_{VV}|^2 + 2|s_{HV}|^2 + |s_{HH}|^2 = |\lambda_1|^2 + |\lambda_2|^2 \tag{80}$$

Both parameters are invariant regardless of polarization basis.

If the power polarization anisotropy coefficient is equal to 0, the target is isotropic target (sphere, plate, etc.). If the parameter is close to 1, the target is anisotropic target (wire, cylinder, etc.).

Figure 41 shows distribution of power polarization anisotropy coefficient around target (5 × 11 pixels, complete polarimetric image consists of 64 × 150 pixels).

The average power polarization anisotropy coefficient is 0.21, Fig. 41a, and 0.88, Fig. 41b, respectively. These results indicate eloquently that the plate behaves like isotropic target, while, the pipe as anisotropic one.

Thus, polarimetric characteristics of buried metallic objects presented above change significantly by target type (isotropic and anisotropic). This means that radar polarimetry improves the detection capability of buried targets.

In (DeBoo et al. 2005) the Mueller matrix data of different man-made samples (e.g., fabric, concrete, metal) as well as their reduced polarimetric data (e.g., depolarization and diattenuation) are studied as a function of scattering geometry at 808 nm (laser diode Hamamatsu L-8446). The seven man-made samples measured are presented in Table 6.

As an example, Fig. 42 depicts the Mueller matrix elements as a function of sample rotation angle for (a) a green painted metal sample and (b) canvas material.

The Mueller matrix data allows making some common conclusions on polarization properties of the samples considered. The green painted metal and the

Table 6 Man-made samples characterized by scatter polarimetry (after DeBoo et al. 2005)

Sample	Description
Metal box part	Smooth metal painted dark green
Nylon plastic material	Dark green color, lightweight material
Canvas with paint splotches	Varies in color and material, underlying cloth is greenish-khaki
Sidewalk concrete	Flat, rough surface
Glass diffuser	Frosted on only one side
Gold-coated glass diffuser	Similar to glass diffuser, roughened surface has thin coating of gold
Screen mesh	Finely meshed material, as used for a window screen

gold-coated diffuser are characterized by large m_{11} peaks at the origin and, hence, exhibit pronounced specular components. Diffusely scattering samples (the nylon plastic, screen mesh fabric, and concrete) exhibit small specular components yielding nearly flat m_{11} curves. The majority of the nonzero matrix elements for all samples are m_{11}, m_{12}, m_{21}, m_{22}, m_{23}, m_{32}, $m_{33,}$ and m_{44}. Correspondingly, nearly zero elements are m_{13}, m_{14}, m_{24}, m_{31}, m_{34}, m_{41}, m_{42}, and m_{43}. Utilization of polar decomposition Eq. (41) shows that these samples being very different, all have negligible retardance and a small diattenuation and polarizance oriented with the s or p planes, but all exhibit high depolarization.

The depolarization index Eq. (29) versus sample rotation angle is shown for five samples in Fig. 43A.

The depolarization profiles in Fig. 43A with changing sample rotation angle resemble inverted Gaussian curves. However, depolarization profiles of the screen mesh and nylon plastic material, which are depicted in Fig. 43B, are not well fit by Gaussian curves.

For most samples the depolarization separately of each linear state and of both circular states are nearly equal. However, the depolarization difference between linear and circular states is significant; circular states are depolarized more than linear states. This is consistent with previous results (Lewis et al. 1998; Mishchenko and Hovenier 1995). Less difference between the depolarization of circular and linear states is observed for three of the more reflective samples, the green painted metal, and the two diffuser samples.

Measurements of the monostatic bidirectional reflectance distribution function (mBRDF) and the complete Mueller matrix were used in Jones et al. (2006) for separability studies of a number of urban (construction) type materials over a broad spectral region from the ultraviolet to the far infrared.

The instrument used in these experiments is a dual rotating retarder spectropolarimeter, described by Azzam (1997), operating in the monostatic reflection mode, Fig. 44.

This polarimeter consists of a PSG before the sample and a PSA after the sample. The PSG consists of a linear polarizer followed by a quarter wave retarder.

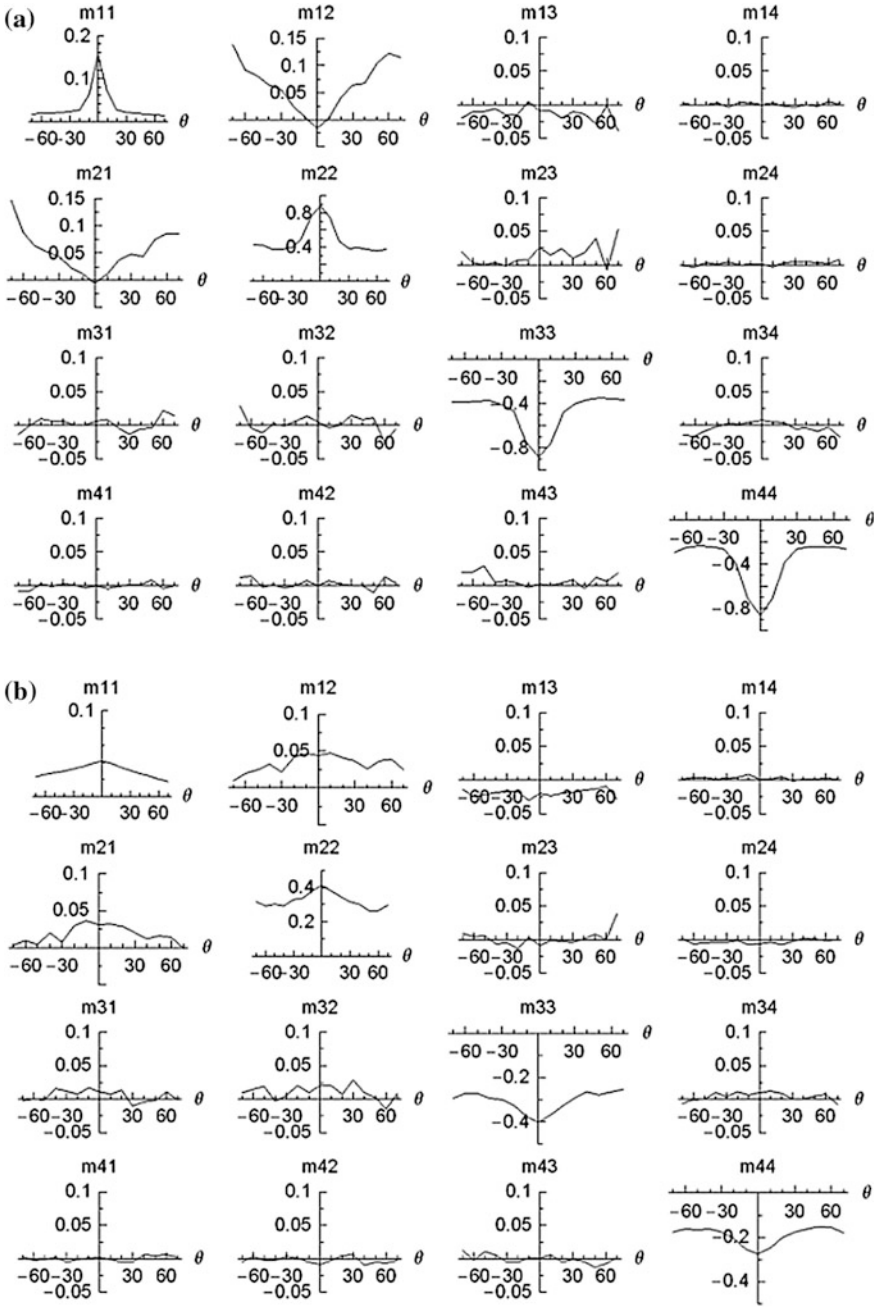


Fig. 42 Normalized Mueller matrix elements for **a** a green-painted metal sample, **b** canvas material as a function of sample rotation angle (after DeBoo et al. 2005)

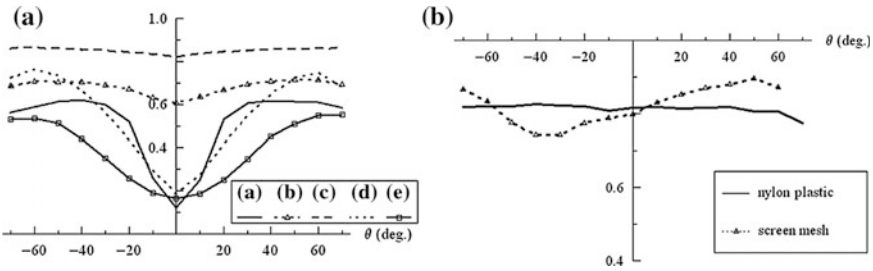


Fig. 43 Depolarization index versus sample rotation angle for **A** samples with Gaussian profiles ((*a*) painted metal; (*b*) canvas; (*c*) concrete; (*d*) glass diffuser; and (*e*) gold-coated diffuser); **B** the screen mesh and nylon plastic material (after DeBoo et al. 2005)

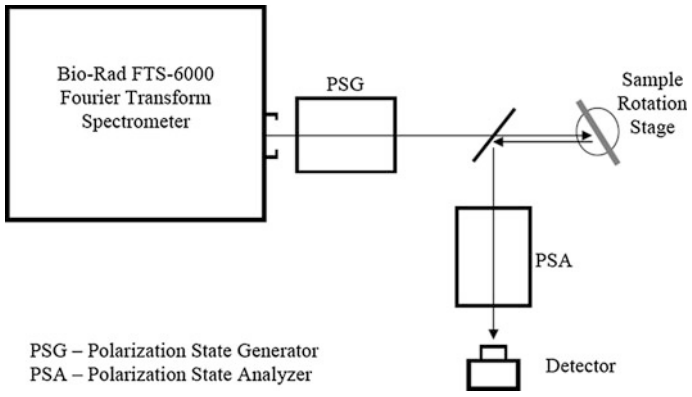


Fig. 44 Optical schematic of the spectropolarimeter (after Jones et al. 2006)

The PSA consists of a quarter wave retarder followed by a linear polarizer and is located on the platter in front of the detector assembly.

Six samples were selected for analysis: rubber, shingle, plywood, drywall (sheetrock) front side, brick, and concrete. Photographs of the test samples are shown in Fig. 45.

For each sample, the mBRDFs and Mueller matrices were measured from 0.7 to 2.3 μm at 1290 distinct frequencies, each separated by wave numbers of approximately 16 cm^{-1}. Each sample is measured at 18 incidence angles (relative to the source) from −10° to +60° with finer resolution near normal incidence (−10°, −8°, −6°, −4°, −2°, −1°, 0°, 1°, 2°, 4°, 6°, 8°, 10°, 15°, 20°, 30°, 45°, 60°). The mBRDFs for all samples are shown in Fig. 46.

As an example, Fig. 47 shows the Mueller matrix data for drywall sample.

The wavelength cross sections of the Mueller matrix data at normal incidence angle for all samples is depicted in Fig. 48.

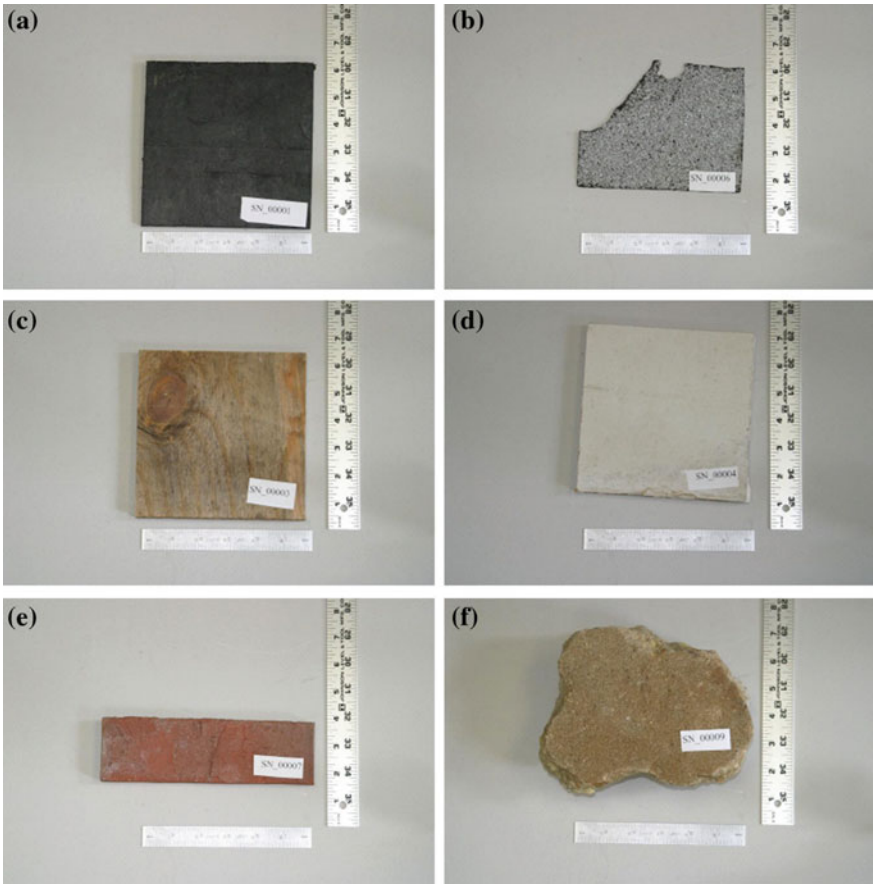


Fig. 45 Material selection: **a** rubber; **b** shingle; **c** plywood; **d** drywall front side; **e** brick; **f** concrete (after Jones et al. 2006)

For solving of the object discrimination problem support vector machine (SVM) with feature vectors containing the three elements: mBRDF (ref), degree of polarization (DoP), and retardance (ret) has been implemented. The retardance is derived by the Lu–Chipman decomposition (Lu and Chipman 1996) of an experimental Mueller matrix. To derive the necessary training and test vectors for the SVM classifiers, authors use the 18 angular measurements at 1.55 μm and create Gaussian statistical models for each feature of each material under consideration. Next, 1000 samples are generated for training and testing: 20% for training, 80% for testing. The results are presented in Fig. 49.

Thus, the support vector machine approach using the statistical models of samples based on reflectance (mBRDF), retardance, and degree of polarization provides adequate separability for the samples under study and yielded a high probability of correct classification while maintaining a low number of false alarms.

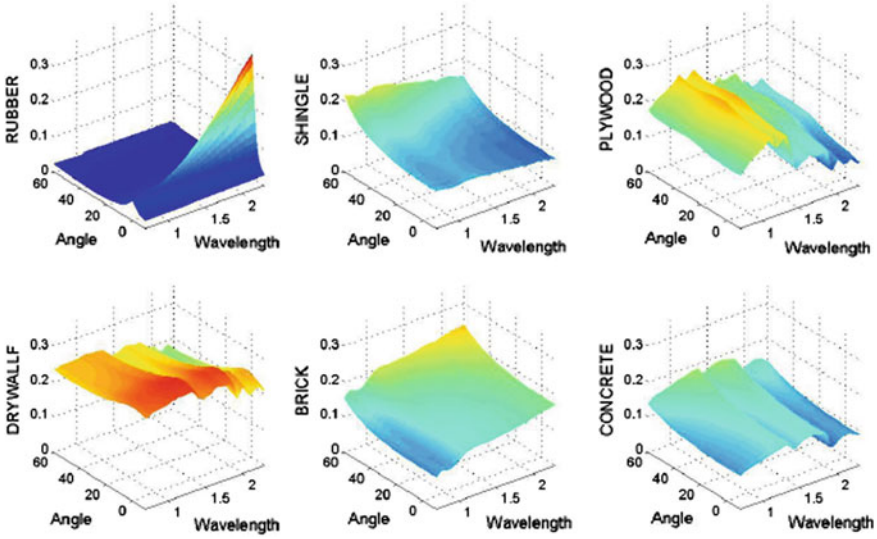


Fig. 46 mBRDFs for all samples (after Jones et al. 2006)

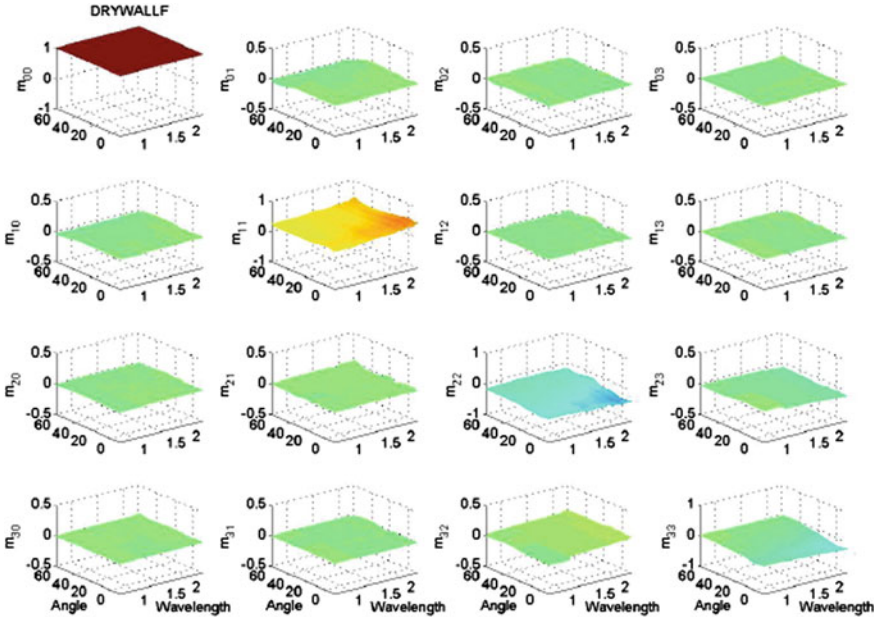


Fig. 47 Mueller matrices for drywall sample (after Jones et al. 2006)

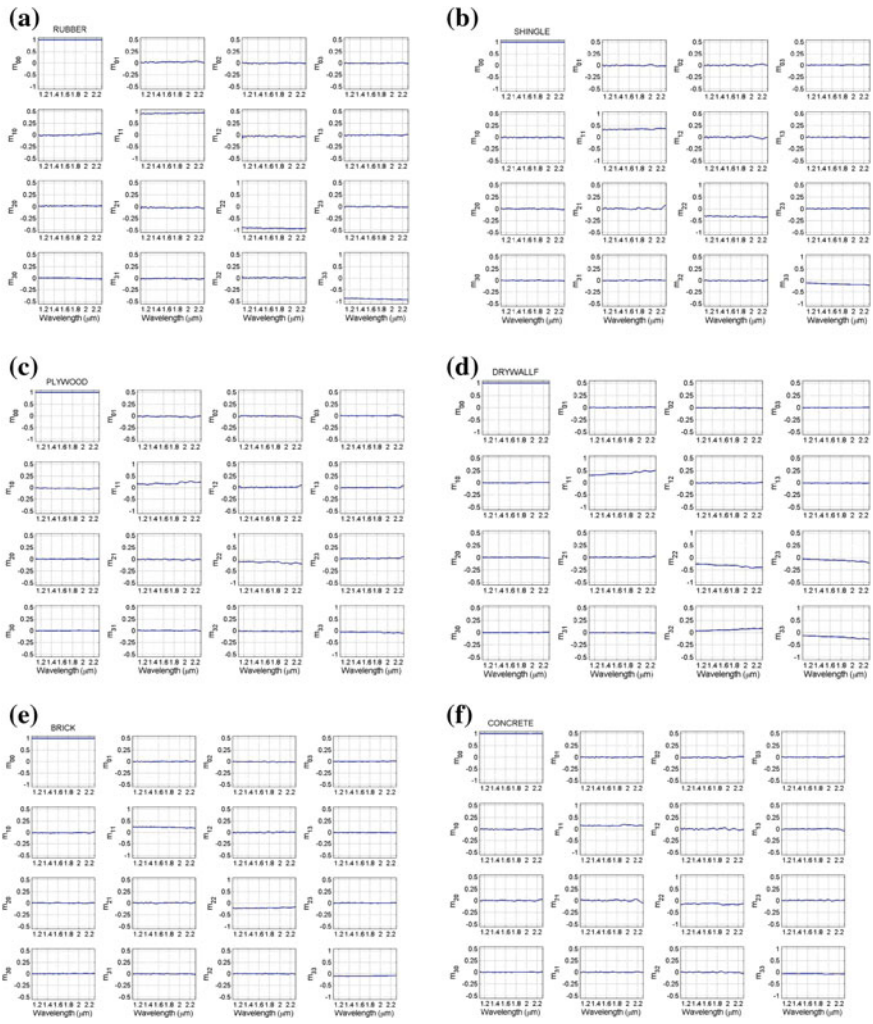


Fig. 48 Mueller matrices at normal incidence in SWIR band for **a** rubber; **b** shingle; **c** plywood; **d** drywall; **e** brick; and **f** concrete (after Jones et al. 2006)

In Noble et al. (2009) the retardance and diattenuation parameters derived from Lu–Chipman Mueller matrix decomposition, see Sect. 3, have been used to determine orientation of a sample with a well-defined texture rotating about the surface normal. The sample is a 4 × 3 in. block of aluminum, which was sanded in one direction with 220 count sandpaper to produce linear grooves typically 2.4 μm deep with aperiodic structure, which was selected, see Fig. 50. In Fig. 50 white regions are areas of low signal where signal-to-noise ratio is poor and where the reflected light is not substantially captured by the instrument due to high slope.

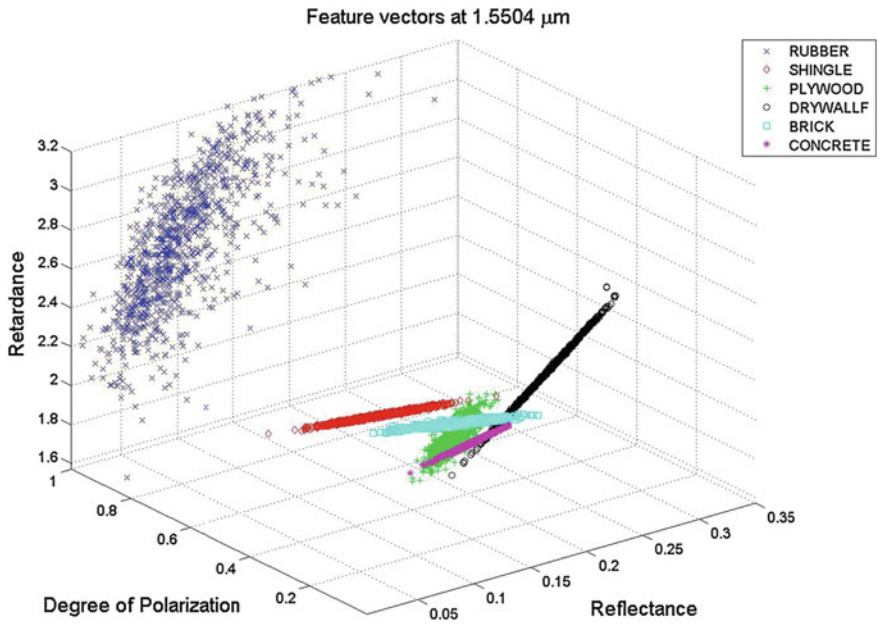


Fig. 49 Feature space for the object discrimination (after Jones et al. 2006)

The specular and diffusely reflected Mueller matrices of sample were measured using dynamic polarimeter operating at a wavelength of 700 nm (Azzam 1997). The beam size at the sample was 1×2 cm. Mueller matrices were acquired at 24 groove orientations as the sample was rotated about its normal from 0° to 360°.

To account for the measurement configuration, Fig. 50, Mueller matrices were multiplied by a reflection matrix of the form

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{81}$$

Figure 51 shows the linear horizontal/vertical, ±45°, and circular values of diattenuation and retardance derived from measured normalized Mueller matrices for the sanded aluminum as a function of the mean groove orientation angle for the specular and off-specular measurement schemes.

All above dependences demonstrate a nearly sinusoidal variation with groove orientation. It can be seen that the diattenuation parameters are more effective.

Authors concluded that the method is suitable for samples with more complex surfaces than those discussed above. However, exact types of rough surfaces for which the method can be used require additional examination.

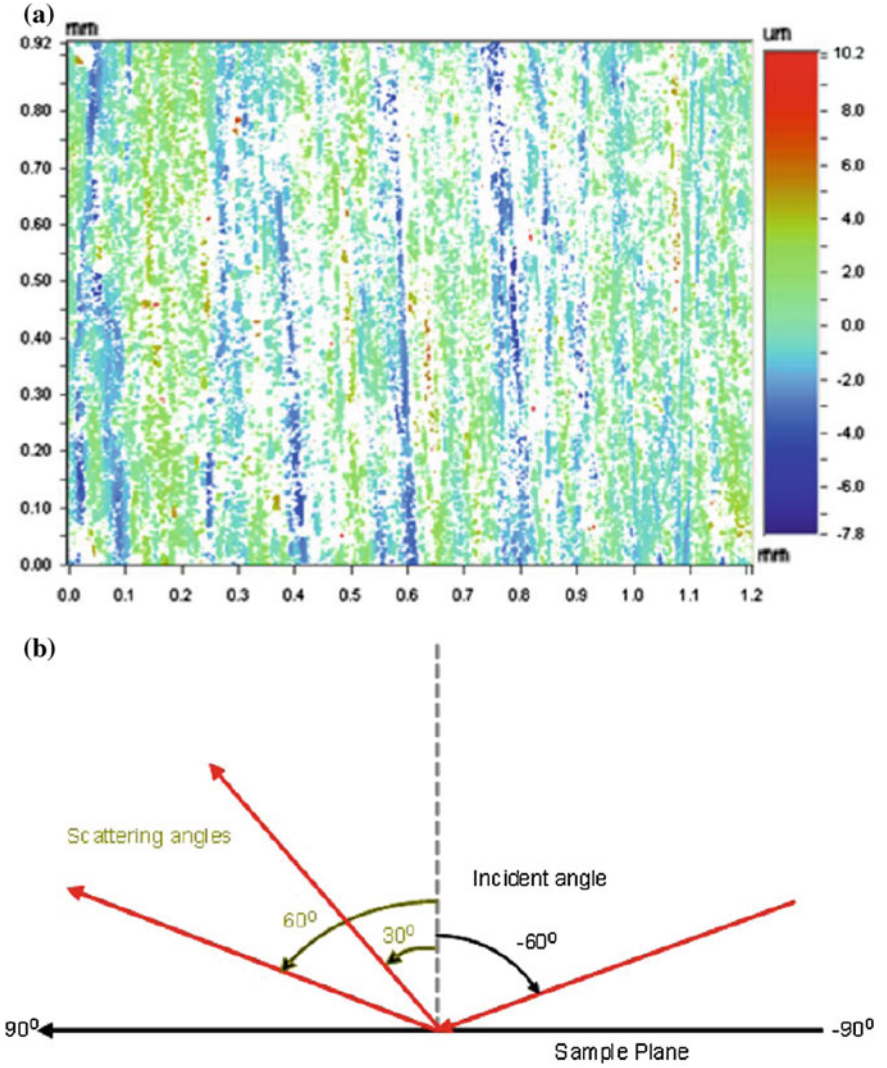


Fig. 50 **a** Surface profile of sanded aluminum sample (heights are in nm). **b** Schemes of the experiment: specular configuration with −60° incident, 60° scattering and off-specular configuration with −60° incident, 30° scattering (after Noble et al. 2009)

The data on polarimetric characteristics, anisotropy and depolarization, of the wood sample (plate) obtained using Lu–Chipman decomposition, Sect. 3, were reported in Srinivasa Reddy et al. (2010).

Experimental setup for measuring Mueller matrix elements and wood sample (*Wrigtia tinctoria*) is presented in Fig. 52. As light source a HeNe laser (Suresh Indu laser, India, 632.8 nm, 20 mW) was used.

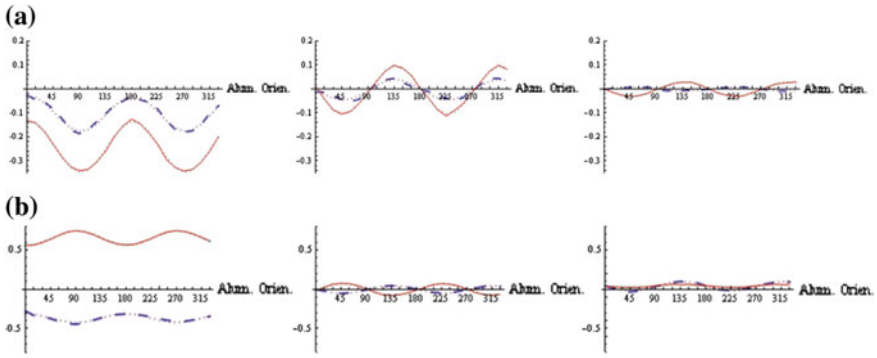


Fig. 51 Diattenuation (**a**) and retardance (**b**) parameters with fit (*dotted lines*) for the specular angle (*solid line*) and the off-specular angle (*dashed line*) (after Noble et al. 2009)

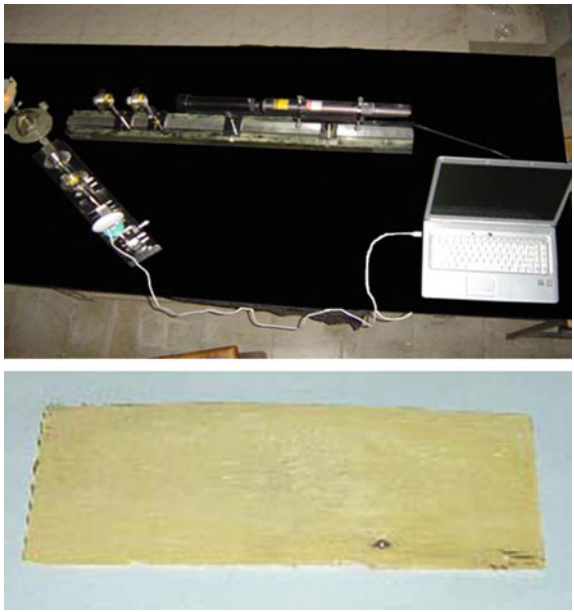


Fig. 52 Schematic experimental setup for obtaining Mueller matrix elements and wood sample (*Wrigtia tinctoria*) (after Srinivasa Reddy et al. 2010)

The sample under consideration is a wood plate obtained from Wrightia Tinctoria, belonging to Apocynaceae family, which has a common name Pala Indigo Plant, Dyers's Oleander. This plant plays an important role in traditional herbal remedies and is under investigation for antibacterial, antineoplastic, and other pharmaceutical functions.

The material obtained is dried for 1 year in order to evaporate the moisture content naturally without losing the chemical composition. The sample is finely polished to have mean thickness of 1.15 mm and average width and length 58.25

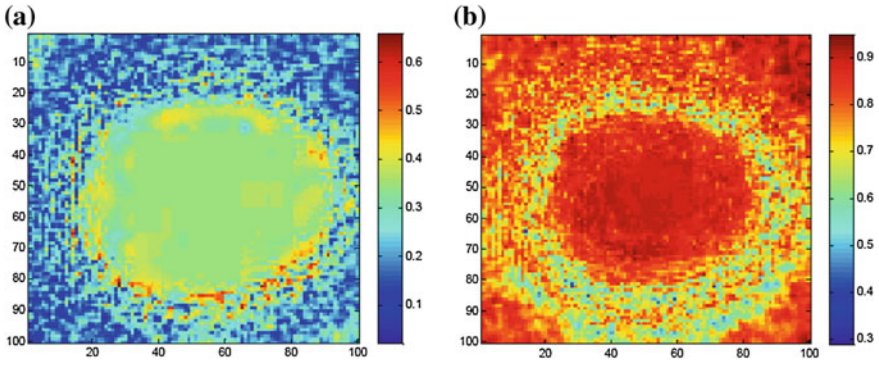


Fig. 53 Diattenuation (**a**) and depolarization (**b**) of the wood sample (after Srinivasa Reddy et al. 2010)

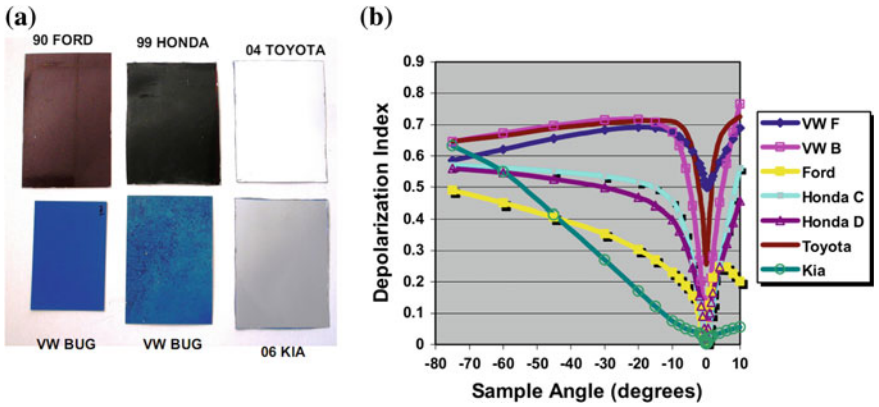


Fig. 54 **a** automobile panel samples and **b** depolarization index versus sample angle at 1.06 μm (Goldstein 2008)

and 136.45 mm, respectively. Figure 53 shows the results for diattenuation (mean value 0.2718) and depolarization (mean value 0.7881) of the sample.

Mueller matrix measurements and depolarization index calculations were used in (Goldstein 2008) for identification of samples of totally different nature, namely, automobile panel samples at the range 0.7–2.6 μm. The samples selected are very shiny: 90 Ford (mean roughness 0.37 μm), 04 Toyota, and 06 Kia; shiny but somewhat dulled: 99 Honda; a satin finish: back of the *VW* bug (mean roughness 0.62 μm); and very aged and rough: front of the *VW* bug (mean roughness 1.66 μm), see Fig. 54a. Depolarization index versus sample angle at 1.06 μm is presented in Fig. 54b.

Again depolarization index features enable to discriminate reliably between automobile panel samples. However, it seems reasonable to use other depolarization

metrics, see Sect. 3, for improving further the discrimination procedure. Thus, depolarization offers an extremely informative metrics for identification of both surface scattering properties such as roughness or reflectivity and bulk scattering with large dynamic range.

References

Abhyankar KD, Fymat AL (1969) Relations between the elements of the phase matrix for scattering. J Math Phys 10:1935–1938

Agrawal AB, Boerner WM (1989) Redevelopment of Kennaugh's target characteristic polarization state theory using the polarization transformation ratio formalism for the coherent case. IEEE Trans Geosci Remote Sens 27:2–14

Anderson DGM, Barakat R (1994) Necessary and sufficient conditions for a Mueller matrix to be derivable from a Jones matrix. J Opt Soc Am A 11:2305–2319

Azzam RM (1997) Mueller-matrix ellipsometry: a review. In: Proceedings of SPIE, vol 3121, pp 396–405

Azzam RM, Bashara NM (1977) Ellipsometry and Polarized Light (North–Holland, New York)

Azzam RM, Bashara NM (1987) Ellipsometry and polarized light. North-Holland, New York

Barakat R (1963) Theory of the coherency matrix for light of arbitrary spectral bandwidth. J Opt Soc Am 53:317–322

Barakat R (1987) Conditions for the physical realizability of polarization matrices characterizing passive systems. J Mod Opt 34:1535–1544

Bicout D, Brosseau Ch, Martinez AS, Schmitt JM (1994) Depolarization of multiply scattered waves by spherical diffusers: influence of the size parameter. Phys Rev E 49:1767–1770

Bjerhammar A (1951) Application of calculus of matrices to method of least squares; with special references to geodetic calculations. Trans Inst Technol (Stockholm) 49

Boerner W-M (1992) Direct and inverse methods in radar polarimetry. Kluwer, Dordrecht

Boerner WM, Mott H, Luneburg E, Livingstone C, Brisco B, Brown RJ, Paterson JS, Cloude SR, Krogager E, Lee JS, Schuler DL, van Zyl JJ, Randall D, Budkewitsch P, Pottier E (1998) Polarimetry in radar remote sensing: basic and applied concepts. In: Manual of remote sensing: principles and applications of imaging radar. Wiley, New York

Bohren CF, Huffman ER (1983) Absorption and scattering of light by small particles. Wiley, New York

Borgeaud M, Shin RT, Kong JA (1987) Theoretical models for polarimetric radar clutter. J. Electromagn Waves Appl 1:73–89

Brosseau Ch (1990) Analysis of experimental data for Mueller polarization matrices. Optik 85:83–86

Brosseau Ch (1998) Fundamentals of polarized light: a statistical optics approach. Wiley, New York

Brosseau C, Givens CR, Kostinski AB (1993) Generalized trace condition on the Mueller-Jones polarization matrix. J Opt Soc Am A 10:2248–2251

Brown CE, Fingas MF (2003) Synthetic aperture radar sensors: viable for marine oil spill response? In: Proceedings of 26th AMOP, pp 299–310

Bueno JM (2001) Indices of linear polarization for an optical system. J Opt A: Pure Appl Opt 3:470–476

Carrieri AH (1999) Neural network pattern recognition by means of differential absorption Mueller matrix spectroscopy. Appl Opt 38:3759–3766

Carrieri AH, Bottiger JR, Owens DJ, Roese ES (1998) Differential absorption Mueller matrix spectroscopy and the infrared detection of crystalline organics. Appl Opt 37:6550–6557

Carrieri AH, Owens DJ, Roese ES, Hung KC, Lim PI, Schultz JC, Talbard MV (2007) Photopolarimetric lidar dual-beam switching device and Mueller matrix standoff detection method. J Appl Remote Sens 1:013502

Carrieri AH, Owens DJ, Schultz JC (2008) Infrared Mueller matrix acquisition and preprocessing system. Appl Opt 47:5019–5027

Carrieri AH, Copper J, Owens DJ, Roese ES, Bottiger JR, Everly RD, Hung KC (2010) Infrared differential-absorption Mueller matrix spectroscopy and neural network-based data fusion for biological aerosol standoff detection. Appl Opt 49:382–393

Chipman RA (1995) Polarimetry. In: Handbook of optics, vol II. McGraw Hill, New York

Chipman RA (1999) Depolarization. In: Proceedings of SPIE, vol 3754, pp 14–20

Chipman RA (2005) Depolarization index and the average degree of polarization. Appl Opt 44:2490–2495

Cloude SR (1986) Group theory and polarization algebra. Optik 7:26–36

Cloude SR (2010) Polarization: applications in remote sensing. Oxford University Press, New York

Cloude SR, Pottier E (1995) Concept of polarization entropy in optical scattering. Opt Eng 34:1599–1610

Cloude SR, Pottier E (1996) A review of target decomposition theorems in radar polarimetry. IEEE Trans Geosci Remote Sens 34:498–518

Cloude SR, Pottier E (1997) An entropy based classification scheme for land applications of polarimetric SAR. IEEE Trans Geosci Remote Sens 35:68–78

Cloude SR, Papathanassiou KP, Pottier E (2001) Radar polarimetry and polarimetric interferometry. IEICE Trans Electron, E84-C:1814–1822

Collett E (1993) Polarized light: fundamentals and applications. Marcel Decker, New York

DeBoo BJ, Sasian JM, Chipman RA (2005) Depolarization of diffusely reflecting man-made objects. Appl Opt 44:5434–5445

Deibler LL, Smith MH (2001) Measurement of the complex refractive index of isotropic materials with Mueller matrix polarimetry. Appl Opt 40:3659–3667

De Martino A, Kim Y-K, Garcia-Caurel E, Laude B and Drevillon B (2003) Optimized Mueller polarimeter with liquid crystals. Opt Lett 28: 616–618

Demirev PA, Feldman AB, Lin JS (2005) Chemical and biological weapons: current concepts for future defenses. Johns Hopkins APL Tech Dig 26:321–333

Dong Y, Richards JA (1995a) Studies of the cylinder-ground double bounce scattering mechanism in forest backscatter models. IEEE Trans Geosci Remote Sens 33:229–231

Dong Y, Richards JA (1995b) Forest discrimination using SAR multifrequency and multipolarization data. Proc IGARSS

Dong Y, Forster BC, Ticehurst C (1996) Decompositions of radar polarization signatures from built and natural targets. Int Arch Photogrammetry Remote Sens XXXI(Part B7):196–202

Dore N, Patruno J, Sarti F, Ruescas AB, Hernandez M (2010) Monitoring of UNESCO sites in danger, including Samarra (Iraq) and Chan Chan (Peru), using optical and radar data. In: Proceedings of EARSeL

Dore N, Patruno J, Pottier E, Crespi M (2013) New research in polarimetric SAR technique for archaeological purposes using ALOS PALSAR data. Archaeol Prospect 20:79–87

Dubois PC, Norikane I (1987) Data volume reduction for imaging radar polarimetry. Proc IGARSS 691–696

Elachi C (1987) Introduction to the physics and techniques of remote sensing. Wiley, New York

Espinosa-Luna R, Bernabeu E (2007) On the Q(M) depolarization metric. Opt Commun 277:256–258

Evans DL, Farr TG, van Zyl JJ, Zebker HA (1988) Radar polarimetry: analysis tools and applications. IEEE Trans Geosci Remote Sens 26:774–789

Freeman A, Durden SL (1998) A three-component scattering model for polarimetric SAR data. IEEE Trans Geosci Remote Sens 36:345–351

Fry ES, Kattawar GW (1981) Relationships between elements of the Stokes matrix. Appl Opt 20:2811–2814

Gade M, Alpers W, Huhnerfuss H, Masuko H, Kobayashi T (1998) Imaging of biogenic and anthropogenic ocean surface films by the multifrequency/multipolarization SIR-C/X-SAR. J Geophys Res 103:18851–18866

Gerrard A, Burch JM (1975) Introduction to matrix methods in optics. Wiley, New York

Gil JJ (2000) Characteristic properties of Mueller matrices. J Opt Soc Am A 17:328–334

Gil JJ (2007) Polarimetric characterization of light and media. Eur Phys J Appl Phys 40:1–47

Gil JJ, Bernabeu E (1985) A depolarization criterion in Mueller matrices. Opt Acta 32:259–261

Gil JJ, Bernabeu E (1986) Depolarization and polarization indexes of an optical system. Opt Acta 33:185–189

Gil JJ, Bernabeu E (1987) Obtainment of the polarizing and retardation parameters of a non-depolarizing optical system from the polar decomposition of its Mueller matrix. Optik 76:67–71

Givens CR, Kostinski AB (1993) A simple necessary and sufficient condition on physically realizable Mueller matrices. J Mod Opt 40:471–481

Goldstein DH (2008) Polarization measurements of automobile paints. Proc SPIE 6972:69720V–1

Guissard A (1994) Mueller and Kennaugh matrices in radar polarimetry. IEEE Trans Geosci Remote Sens 32:590–597

Hauge PS (1980) Recent developments in instrumentation in ellipsometry. Surf Sci 96:108–140

Haugland SM, Bahar E, Carrieri AH (1992) Identification of contaminant coatings over rough surfaces using polarized infrared scattering. Appl Opt 31:3847–3852

Hellmann M, Cloude SR (2001) Discrimination between low metal content mine and non-mine-targets using polarimetric ultra-wide-band radar. Int Geosci Remote Sens Symp 3:1113–1115

Hovenier JW (1969) Symmetry relations for scattering of polarized light in a slab of randomly oriented particles. J Atmos Sci 26:488–499

Hovenier JW (1970) Principles of symmetry for polarization studies of planets. Astron Astrophys 7:86–90

Hovenier JW (1994) Structure of a general pure Mueller matrix. Appl Opt 33:8318–8324

Hovenier JW, Mackowski DW (1998) Symmetry relations for forward and backward scattering by randomly oriented particles. J Quant Spectrosc Radiat Transfer 60:483–492

Hovenier JW, van der Mee CVM (2000) Basic relationships for matrices describing scattering by small particles. In: Mishchenko MI, Hovenier JW, Travis LD (eds) Light scattering by nonspherical particles: theory, measurements, and applications. Academic Press, San Diego, pp 61–85

Hovenier JW, van de Hulst HC, van der Mee CVM (1986) Conditions for the elements of the scattering matrix. Astron Astrophys 157:301–310

Hu Ch-R, Kattawar GW, Parkin ME, Herb P (1987) Symmetry theorems on the forward and backward scattering Mueller matrices for light scattering from a nonspherical dielectric scatterer. Appl Opt 26:4159–4173

Hurwitz H, Jones CR (1941) A new calculus for the treatment of optical systems. II. Proof of three general equivalence theorems. J Opt Soc Am 31:493–499

Huynen JR (1970) Phenomenological theory of radar targets. PhD dissertation, Delft Technical University, The Netherlands

Jones RC (1941) A new calculus for the treatment of optical systems. I. Description and discussion of the calculus. J Opt Soc Am 31:488–493

Jones RC (1942) A new calculus for the treatment of optical systems. IV. J Opt Soc Am 32:486–493

Jones RC (1947) A new calculus for the treatment of optical systems. V. A more general formulation and description of another calculus. J Opt Soc Am 37:107–110

Jones RC (1956) A new calculus for the treatment of optical systems. VIII. Electromagnetic theory. J Opt Soc Am 46:126–131

Jones DG, Goldstein DH, Spaulding JC (2006) Reflective and polarimetric characteristics of urban materials. Proc SPIE 6240:62400A(10)

Kaplan B, Ledanois G, Drevillon B (2001) Mueller matrix of dense polystyrene latex sphere suspensions: measurements and Monte Carlo simulation. Appl Opt 40:2769–2777
Kennaugh K (1951) Effects of type of polarization on echo characteristics. Tech. Rep OH 389-4, 35, and OH 381-9, 39
Kennaugh K (1952) Polarization properties of radar reflections. M.Sc. Thesis, The Ohio State University, Columbus, Ohio
Kim K, Mandel L, Wolf E (1987) Relationship between Jones and Mueller matrices for random media. J Opt Soc Am A 4:433–437
Kim YL, Pradhan P, Kim MH, Backman V (2006) Circular polarization memory effect in low-coherence enhanced backscattering of light. Opt Lett 31:2744–2746
Kliger DS, Lewis JW, Randall CE (1990) Polarized light in optics and spectroscopy. Academic Press, New York
Kokhanovsky AA (2001) Light scattering media optics: problems and solutions. Praxis, Chichester, UK
Kokhanovsky AA (2003a) Parameterization of the Mueller matrix of oceanic waters. J Geophys Res 108:3175
Kokhanovsky AA (2003b) Polarization optics of random media. Praxis, Chichester, UK
Kokhanovsky AA (2003c) Optical properties of irregularly shaped particles. J Phys D Appl Phys 36:915–923
Kong JA, Swartz AS, Yueh HA, Novak LM, Shin RT (1987) Identification of terrain cover using the optimum polarimeter classifier. J Electromagn Waves Appl 2:171–194
Kostinski AB (1992) Depolarization criterion for incoherent scattering. Appl Opt 31:3506–3508
Kostinski AB, Givens CR, Kwiatkowski JM (1993) Constraints on Mueller matrices of polarization optics. Appl Opt 32:1646–1651
Krogager E (1990) A new decomposition of the radar target scattering matrix. Electron Lett 26:1525–1526
Lankaster P, Tismenetsky M (1985) The theory of matrices. Academic Press, San Diego
Lasaponara R, Masini N (2013) satellite synthetic aperture radar in archaeology and cultural landscape: an overview. Archaeol Prospect 20:71–78
Lawless R, Xie Y, Yang P et al (2006) Polarization and effective Mueller matrix for multiple scattering of light by nonspherical ice crystals. Opt Express 14:6381–6393
Lee J-S, Pottier E (2009) Polarimetric radar imaging: from basics to applications. CRC Press, New York
Lee J-S, Krogager E, Ainsworth ThL, Boerner W-M (2006) Polarimetric analysis of radar signature of a manmade structure. IEEE Geosci Remote Sens Lett 3:555–559
Lewis GD, Jordan DL, Jakeman E (1998) Backscatter linear and circular polarization analysis of roughened aluminum. Appl Opt 37:5985–5992
Lu S-Y, Chipman RA (1994) Homogeneous and inhomogeneous Jones matrices. J Opt Soc Am A 11:766–773
Lu S-Y, Chipman RA (1996) Interpretation of Mueller matrices based on polar decomposition. J Opt Soc Am A 13:1106–1113
Lueneburg E (1995) Principles of radar polarimetry-using the directional Jones vector approach. IEICE Trans Electron E78-C:1339–1345
Manickavasagam S, Menguc MP (1998) Scattering-matrix elements of coated infinite-length cylinders. Appl Opt 37:2473–2482
Mar'enko VV, Savenkov SN (1994) Representation of arbitrary Mueller matrix in the basis of matrices of circular and linear anisotropy. Opt Spectrosc 76:94–96
Mattia F, Le Toan T, Souyris JC, De Carolis G et al (1997) The effect of surface roughness on multifrequency polarimetric SAR data. IEEE Trans Geosci Remote Sens 35:954–965
McNairna H, Duguay C, Brisco B, Pultz TJ (2002) The effect of soil and crop residue characteristics on polarimetric radar response. Remote Sens Environ 80:308–320
Migliaccio M, Gambardella A, Tranfaglia M (2007) SAR polarimetry to observe oil spills. IEEE Trans. Geosci. Remote Sens. 45:506–511

Migliaccio M, Nunziata F, Gambardella A (2008) Polarimetric signature for oil spill observation. In: Proceedings of US/EU-Baltic international symposium, pp 1–5
Migliaccio M, Nunziata F, Gambardella A (2009a) On the copolarised phase difference for oil spill observation. Int J Remote Sens 30:1587–1602
Migliaccio M, Gambardella A, Nunziata F, Shimada M, Isoguchi O (2009b) The PALSAR polarimetric mode for sea oil slick observation. IEEE Trans Geosci Remote Sens 47:4032–4041
Mishchenko MI, Hovenier JW (1995) Depolarization of light backscattered by randomly oriented nonspherical particles. Opt Lett 20:1356–1358
Mishchenko MI, Travis LD (2000) Polarization and depolarization of light. In: Moreno F, González F (eds) Light scattering from microstructures. Springer, Berlin, pp 159–175
Mishchenko MI, Hovenier JW, Travis LD (eds) (2000) Light scattering by nonspherical particles: theory, measurements, and applications. Academic Press, San Diego
Mishchenko MI, Travis LD, Lacis AA (2002) Scattering, absorption, and emission of light by small particles. Cambridge University Press, Cambridge, UK
Mishchenko MI, Travis LD, Lacis AA (2006) Multiple scattering of light by particles. Radiative transfer and coherent backscattering, Cambridge University Press, Cambridge
Mishchenko MI, Rosenbush VK, Kiselev NN, Lupishko DF, Tishkovets VP, Kaydash VG, Belskaya IN, Efimov YS, Shakhovskoy NM (2010) Polarimetric remote sensing of solar system objects. Akademperiodyka, Kyiv
Moore EH (1920) On the reciprocal of the general algebraic matrix. Bull Am Math Soc 26:394–395
Moriyama T, Nakamura M, Yamaguchi Y, Yamada H, Boerner W-M (1997) Radar polarimetry applied to the classification of target buried in the underground. Proc SPIE 3120:182–189
Moriyama T, Yamaguchi y, Uratsuka s, Umehara t et al (2005) A study on polarimetric correlation coefficient for feature extraction of polarimetric SAR data. IEICE Trans Commun E88-B (6):2355–2361
Mott H (2007) Remote sensing with polarimetric radar. Wiley, New York
Mueller H (1948) The foundation of optics. J Opt Soc Am 38:661
Munoz O, Volten H, de Haan JF, Vassen W, Hovenier JW (2001) Experimental determination of scattering matrices of fly ash and clay particles at 442 and 633 nm. J Geophys Res 106:22833–22844
Munoz O, Volten H, de Haan JF, Vassen W, Hovenier JW (2002) Experimental determination of the phase function and degree of linear polarization of El Chichon and Pinatubo volcanic ashes. J Geophys Res 107(D13):4174
Munoz O, Volten H, Hovenier JW, Veihelmann B, van der Zande WJ, Waters LBFM, Rose WI (2004) Scattering matrices of volcanic ash particles of Mount St. Helens, Redoubt, and Mount Spurr Volcanoes. J Geophys Res 109:D 16201
Muttiah RS (ed) (2002) From laboratory spectroscopy to remotely sensed spectra of terrestrial ecosystems. Kluwer, Dordrecht
Nagirner DI (1993) Constrains on matrices transforming Stokes vectors. Astron Astrophys 275:318–324
Nesti G (1998) Backscattering from rough dielectric surfaces. EMSL Report, SAI-JRC, Ispra, Italy
Noble H, Lam W-ST, Chipman RA (2009) Inferring the orientation of texture from polarization parameters. Proc SPIE 7461:746109(6)
Novak LM, Burl MC (1990) Optimal speckle reduction in polarimetric SAR imagery. IEEE Trans Aerosp Electron Syst 26:293–305
Nunziata F, Gambardella A, Migliaccio M (2008) On the Mueller scattering matrix for SAR sea oil slick observation. IEEE Geosci Remote Sens Lett 5:691–695
Oberemok EA, Savenkov SN (2003) Structure of deterministic Mueller matrices and their reconstruction in the method of three input polarizations. J Appl Spectrosc 70:224–229
Ossikovski R (2008) Interpretation of nondepolarizing Mueller matrices based on singular-value decomposition. J Opt Soc Am A 25:473–482

Ossikovski R (2009) Analysis of depolarizing Mueller matrices through a symmetric decomposition. J Opt Soc Am A 26:1109–1118
Park S-E, Yamaguchi Y, Kim D-J (2013) Polarimetric SAR remote sensing of the 2011 Tohoku earthquake using ALOS/PALSAR. Remote Sens Environ 132:212–220
Parke NG III (1949) Optical algebra. J Math Phys 28:131–139
Parke NG III (1948) Matrix optics. PhD thesis, Massachusetts Institute of Technology, Cambridge
Patruno J, Dore N, Crespi M, Pottier E (2013) Polarimetric multifrequency and multi-incidence SAR sensors analysis for archaeological purposes. Archaeol Prospect 20:89–96
Penrose R (1955) A generalized inverse for matrices. Proc Camb Philos Soc 151:406–413
Perrin F (1942) Polarization of light scattering by isotropic opalescent media. J Chem Phys 10:415–427
Petryk MWP (2007) Promising spectroscopic techniques for the portable detection of condensed-phase contaminants on surfaces. Appl Sp Rev 42:287–343
Potton RJ (2004) Reciprocity in optics. Rep Prog Phys 67:717–754
Priezzhev AV, Tuchin VV, Shubochkin LP (1989) Laser diagnostics in biology and medicine. Nauka, Moscow (in Russian)
Rao AVG, Mallesh KS, Sudha (1998a) On the algebraic characterization of a Mueller matrix in polarization optics. I. Identifying a Mueller matrix from its N matrix. J Mod Opt 45:955–987
Rao AVG, Mallesh KS, Sudha (1998b) On the algebraic characterization of a Mueller matrix in polarization optics. II. Necessary and sufficient conditions for Jones-derived Mueller matrices. J Mod Opt 45:989–999
Ray TW, Farr TG, van Zyl JJ (1992) Detection of land degradation with polarimetric SAR. Geophys Res Lett 19:1587–1590
Rojas-Ochoa LF, Lacoste D, Lenke R et al (2004) Depolarization of backscattered linearly polarized light. J Opt Soc Am A 21:1799–1804
Savenkov SN (2002) Optimization and structuring of the instrument matrix for polarimetric measurements. Opt Eng 41:965–972
Savenkov SN (2007) Analysis of generalized polarimetric measurement equation. Proc SPIE 6682:668214
Savenkov SN, Muttiah RS (2004) Inverse polarimetry and light scattering from leaves. In: Photopolarimetry in remote sensing. Springer, Netherlands
Savenkov SN, Muttiah RS, Oberemok YA (2003) Transmitted and reflected scattering matrices from an English oak leaf. Appl Opt 42:4955–4962
Savenkov SN, Sydoruk OI, Muttiah RS (2005) The conditions for polarization elements to be dichroic and birefringent. J Opt Soc Am A 22:1447–1452
Savenkov SN, Marienko VV, Oberemok EA, Sydoruk OI (2006) Generalized matrix equivalence theorem for polarization theory. Phys Rev E 74:056607
Savenkov SN, Muttiah RS, Yushtin KE, Volchkov SA (2007a) Mueller-matrix model of an inhomogeneous, linear, birefringent medium: single scattering case. J Quant Spectrosc Radiat Transfer 106:475–486
Savenkov SN, Sydoruk OI, Muttiah RS (2007b) Eigenanalysis of dichroic, birefringent, and degenerate polarization elements: a Jones-calculus study. Appl Opt 46:6700–6709
Saxon DS (1955) Tensor scattering matrix for electromagnetic fields. Phys Rev 100:1771–1775
Schuler DL, Lee JS, De Grandi G (1996) Measurement of topography using polarimetric SAR images. IEEE Trans Geosci Remote Sens 34:1266–1277
Sekera Z (1966) Scattering matrices and reciprocity relationships for various representations of the state of polarization. J Opt Soc Am 56:1732–1740
Shindo Y (1995) Applications of polarized modulator technique in polymer science. Opt Eng 34:3369–3384
Shurcliff WA (1962) Polarized light: production and use. Harvard University Press, Harvard
Simon R (1982) The connection between Mueller and Jones matrices of polarization optics. Opt Commun 42:293–297
Simon R (1987) Mueller matrices and depolarization criteria. J Mod Opt 34:569–575

Sinclair G (1950) Transmission and reception of elliptically polarized waves. Proc IRE 38:148–151
Smith MH (2002) Optimization of a dual-rotating-retarder Mueller matrix polarimeter. Appl Opt 41:2488–2495
Sneh A, Weissbrod T (1973) Nile delta: the defunct Pelusiac branch identified. Science 180:59–61
Solleillet P (1929) Sur les paramètres caractérisant la polarisation partielle de la lumière dans les phénomènes de fluorescence. Ann Phys Biol Med 12:23–97
Sridhar R, Simon R (1994) Normal form for Mueller matrices in polarization optics. J Mod Opt 41:1903–1915
Srinivasa Reddy K, Mohan kumar V, Chandralingam S, Raghavendra Rao P, Kanaka Rao PV (2010) Optical signature of wood sample—Mueller matrix imaging polarimetry. ARPN J Eng Appl Sci 5:34–38
Stabo-Eeg F, Letalick D, Steinvall O, Lindgren M (2008a) Discriminating land mines from natural backgrounds by depolarization. Proc SPIE, 7114:71140H (10)
Stabo-Eeg F, Nerbo IS, Kildemo M, Lindgren M (2008b) A well conditioned multiple laser Mueller matrix ellipsometer. Opt Eng 47(7):073604
Stabo-Eeg F, Kildemo M, Garcia-Caurel E, Lindgren M (2008c) Design and characterization of achromatic 132° retarders in CaF_2 and fused silica. J Mod Opt 55:2203–2214
Stewart Ch, Lasaponara R, Schiavon G (2013) ALOS PALSAR analysis of the archaeological site of pelusium. Archaeol Prospect 20:109–116
Swami MK, Manhas S, Buddhiwant P, Ghosh N, Uppal A, Gupta PK (2006) Polar decomposition of 3x3 Mueller matrix: a tool for quantitative tissue polarimetry. Opt Express 14:9324–9337
Tang ST, Kwok HS (2001) 3x3 Matrix for unitary optical systems. J Opt Soc Am A 18:2138–2145
Touzi R, Lopes A (1994) The principle of speckle filtering in polarimetric SAR imagery. IEEE Trans Geosci Remote Sens 32:1110–1114
Touzi R, Boerner WM, Lee JS, Lueneburg E (2004) A review of polarimetry in the context of synthetic aperture radar: concepts and information extraction. Can J Remote Sens 30:380–407
Tuchin VV (ed) (2002) Handbook of optical biomedical diagnostics. SPIE Press, Bellingham
Tuchin VV, Wang LV, Zimnyakov DA (2006) Optical Polarization in biomedical applications. Springer, Berlin
Twietmeyer K, Chipman RA (2008) Optimization of Mueller polarimeters in the presence of error sources. Opt Express 16:11589–11603
Tyo JS, Goldstein DL, Chenault DB, Shaw JA (2006) Review of passive imaging polarimetry for remotesensing applications. Appl Opt 45:5453–5469
Tyo JS, Wang Zh, Johnson SJ, Hoover BG (2010) Design and optimization of partial Mueller matrix polarimeters. Appl Opt 49:2326–2333
Ulaby FT, Moore RK and Fung AK (1982) Microwave Remote Sensing Active and Passive. Vol. II Radar Remote Sensing and Surface Scattering and Emission Theory (Addison–Wesley, New York)
Valenzuela GR (1967) Depolarization of EM waves by slightly rough surfaces. IEEE Trans Antennas Propag AP-15(4):552–557
van de Hulst HC (1957) Light scattering by small particles. Wiley, New York
van der Mee CVM (1993) An eigenvalue criterion for matrices transforming Stokes parameters. J Math Phys 34:5072–5088
van der Mee CVM, Hovenier JW (1992) Structure of matrices transforming Stokes parameters. J Math Phys 33:3574–3584
van Zyl JJ (1989) Unsupervised classification of scattering behavior using radar polarimetry data. IEEE Trans Geosci Remote Sens 27:37–45
van Zyl JJ (1992) Application of Cloude's target decomposition theorem to polarimetric imaging radar data. In: Proceedings of SPIE, vol 1748, pp 184–191
van Zyl JJ, Ulaby FT (1990) Scattering matrix representation for simple targets. In: Radar polarimetry for geosciences applications. Artech House, Norwood
van Zyl JJ, Zebker HA (1990) Imaging radar polarimetry. In: PIER 3, progress in electromagnetics research. Elsevier, New York

van Zyl JJ, Zebker HA, Elachi C (1987) Imaging radar polarization signatures: theory and observation. Radio Sci 22:529–543

Vansteenkiste N, Nignolo P, Aspect A (1993) Optical reversibility theorems for polarization: application to remote control of polarization. J Opt Soc Am A 10:2240–2245

Volten H, Munoz O, Rol E, de Haan JF, Vassen W, Hovenier JW, Muinonen K, Nousianen T (2001) Scattering matrices of mineral aerosol particles at 441.6 nm and 632.8 nm. J Geophys Res 106(17):375

Voss KJ, Fry ES (1984) Measurement of the Mueller matrix for ocean water. Appl Opt 23:4427–4436

Whitney C (1971) Pauli-algebraic operators in polarization optics. J Opt Soc Am 61:1207–1213

Woodhouse IH (2006) Predicting backscatter-biomass and height-biomass trends using a macroecology model. IEEE Trans Geosci Remote Sens GRS-44:871–877

Xing Z-F (1992) On the deterministic and nondeterministic Mueller matrix. J Mod Opt 39:461–484

Xu L, Ding J, Cheng AYS (2002) Scattering matrix of infrared radiation by ice finite circular cylinders. Appl Opt 41:2333–2348

Yamaguchi Y, Moriyama T, Ishido M, Yamada H (2005) Four-component scattering model for polarimetric SAR image decomposition. IEEE Trans Geosci Remote Sens 43(8):1699–1706

Zebker FA, van Zyl JJ (1991) Imaging radar polarimetry: a review. Proc IEEE 79:1583–1606

Zubko E, Shkuratov YG, Videen G (2004) Coherent backscattering effect for non-zero elements of Mueller matrix of discrete media at different illumination-observation geometries. J Quant Spectrosc Radiat Transfer 89:443–452

Index

© Springer-Verlag Berlin Heidelberg 2016

A. Kokhanovsky (ed.), *Light Scattering Reviews, Volume 11*,
Springer Praxis Books, DOI 10.1007/978-3-662-49538-4