

Fundamentals of  
**POLARIZED  
LIGHT**

A Statistical  
Optics  
Approach

Christian Brosseau

Brosseau

# Comprehensive coverage of light polarization theory and its practical applications in today's cutting-edge technologies

Besides being indispensable to modern investigations into the physical world, light polarization is a fundamental component of several revolutionary technological innovations in such diverse fields as telecommunications, pollution control, and medical diagnostics. Yet there is a conspicuous dearth of texts and professional references providing researchers and engineers with a unified, comprehensive treatment of basic light polarization theory and its applications to current microwave and optical technology. This book fills that gap in the literature.

*Fundamentals of Polarized Light* serves equally well as an advanced text for physics and electrical engineering students and a professional reference for practicing engineers and researchers. It combines a rational, integrated presentation of the theory behind modern applications of light polarization with several demonstrations of current applications. A key feature of the book is that the analysis of polarized light and its interaction with linear optical media is presented from a statistical point of view.

Topics covered include:

- Historical foundations of polarized light
- Classical radiation field theory and Maxwell's equations
- Statistical theory of partial polarization, including a discussion of the thermodynamics of radiation fields
- Propagation of polarized light through linear optical systems
- Polarization transfer matrix methods for describing changes in polarization states that occur during reflection and refraction
- Propagation of partially polarized waves in disordered systems and anisotropic media
- Polarizers, compensators, and other optical components
- Measurements of the Jones and Mueller polarization matrices

**CHRISTIAN BROSSEAU** received a PhD degree in physics from Fourier University, Grenoble, France, in 1989. He was a postdoctoral Research Fellow at Harvard from 1989 to 1990 and then joined Fourier University as an associate research physicist. In 1997, he became a physics professor in the Department of Physics at the University of Brest, France. His primary research interests include the physics of heterogeneous materials and theoretical optics.

Cover Design: Anne Michele Abbott

Cover Photograph: Claude Barbier, "American Soda"

# Fundamentals of **POLARIZED LIGHT**

A Statistical

535.52

B874f

**WILEY-INTERSCIENCE**

John Wiley & Sons, Inc.

Professional, Reference and Trade Group

605 Third Avenue, New York, N.Y. 10158-0012

New York • Chichester • Weinheim

Brisbane • Singapore • Toronto

ISBN 0-471-14302-2

9 780471 143024

INTER-  
SCIENCE

---

# **FUNDAMENTALS OF POLARIZED LIGHT**

**A STATISTICAL OPTICS APPROACH**

---

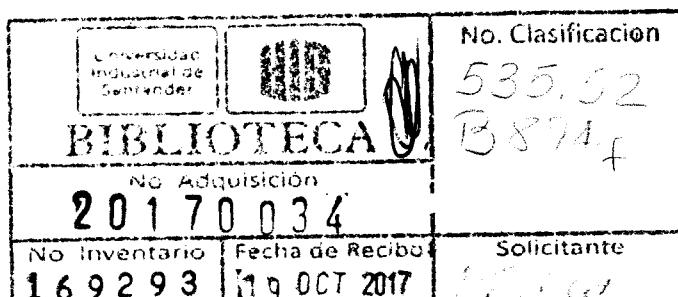
---

# FUNDAMENTALS OF POLARIZED LIGHT

## A STATISTICAL OPTICS APPROACH

---

Christian Brosseau  
University of Brest, France



A WILEY-INTERSCIENCE PUBLICATION

**JOHN WILEY & SONS, INC.**

New York . Chichester . Weinheim . Brisbane . Singapore . Toronto

This book is printed on acid-free paper. ☺

Copyright © 1998 by John Wiley & Sons, Inc. All rights reserved.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923, (508) 750-8400, fax (508) 750-4744. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 605 Third Avenue, New York, NY 10158-0012, (212) 850-6011, fax (212) 850-6008, E-Mail: PERMREQ @ WILEY.COM.

***Library of Congress Cataloging-in-Publication Data:***

Brosseau, Christian, 1961-

Fundamentals of polarized light : a statistical optics approach /  
by Christian Brosseau.

p. cm.

Includes bibliographical references and index.

ISBN 0-471-14302-2 (cloth :alk. paper)

1. Polarization (Light) I. Title.

QC441.B76 1998

535.5'2--dc21

98-10444

## CONTENTS

---

<b>PREFACE</b>	<b>xi</b>
<b>PART 1 HISTORICAL SURVEY OF UNDERSTANDING OF POLARIZED LIGHT</b>	<b>1</b>
<b>1.1. First Period: Early Ideas and Observations—from Bartholinus         to Stokes</b>	<b>3</b>
<b>1.2. Second Period: The Electromagnetic Nature of Light</b>	<b>9</b>
<b>1.3. Third Period: The Coherence and Quantum Properties of Light</b>	<b>15</b>
<b>Further Reading</b>	<b>27</b>
<b>PART 2 PRELIMINARIES TO A CLASSIC RADIATION FIELD THEORY</b>	<b>35</b>
<b>2.1. The Basic Differential Equations and Boundary Conditions</b>	<b>37</b>
<b>2.1.1. Maxwell's Equations</b>	<b>37</b>
<b>2.1.2. Boundary Conditions</b>	<b>41</b>
<b>2.1.3. Fresnel's Formulas for Reflection and Transmission of a             Plane Wave at a Planar Boundary between Homogeneous             Isotropic Media</b>	<b>42</b>
<b>2.1.4. Wiener's Experiment</b>	<b>48</b>
<b>2.2. Invariance Transformations</b>	<b>51</b>
<b>2.2.1. Gauge Transformations</b>	<b>51</b>
<b>2.2.2. Time Inversion</b>	<b>54</b>
<b>2.2.3. Spatial Inversion</b>	<b>55</b>
<b>2.2.4. Duality Transformations</b>	<b>55</b>

<b>2.3. Monochromatic Plane Wave</b>	<b>57</b>
2.3.1. Harmonic Plane-Wave Solution of the Wave Equation	57
2.3.2. Reciprocity Principle	61
2.3.3. Poynting's Theorem	62
2.3.4. Normal Modes Representation of Electromagnetic Field	63
<b>Further Reading</b>	<b>66</b>
<b>PART 3 POLARIZATION AND THE RADIATION FIELD</b>	<b>67</b>
<b>3.1. Elementary Concepts and Definitions</b>	<b>69</b>
3.1.1. Temporal Fluctuations in Light Beams: Characteristic Timescales and Partial Polarization	69
3.1.2. The Analytic Signal Representation	70
3.1.3. The Instantaneous Ellipse for the Electric Field	76
3.1.4. The Complex Polarization Ratio Representation of Polarized Light	84
3.1.5. Statistical Description of Fluctuations of a Partially Polarized Optical Field	88
3.1.5.1. Correlation Functions of Analytic Signals	89
3.1.5.2. The Probability Density Function of the Radiation Field	91
3.1.5.3. Second-Order Coherence Properties of a Stochastic Radiation Field	93
3.1.5.4. Relations between Correlation Tensors and Poynting Vector	95
3.1.5.5. Dynamical Equations for Second-Order Correlation Tensors	96
3.1.5.6. Wiener–Khintchine Theorem and Cross-Spectral Correlation Tensor	97
3.1.5.7. Blackbody Radiation in Equilibrium in an Enclosure	99
3.1.6. Stokes Parameters and Coherency Matrix Formalism	102
3.1.6.1. Stokes Parameters	102
3.1.6.2. Coherency (Density) Matrix Formalism	105
3.1.6.3. Spectral Decomposition Theorem	109
3.1.6.4. Principle of Optical Equivalence	113
3.1.6.5. Reduction Property	115
3.1.6.6. Covariance Matrix of Instantaneous Stokes Parameters	116
3.1.6.7. Digression: Generalized Stokes Parameters	118

<b>3.2. Geometric Representations of Partially Polarized Light</b>	<b>123</b>
3.2.1. The Stokes Vector Space	124
3.2.2. Convexity Structure of Sets of Polarization States	128
3.2.3. Stereographic Projection Plane of Poincaré Sphere	130
3.2.4. Geometric Phase and Polarization Cycles	131
3.2.5. Analogy between Polarization and Two-Level Systems: Bloch Equation	135
<b>3.3. Statistics of the Radiation Field</b>	<b>138</b>
3.3.1. Statistics of Amplitudes and Phases for a Gaussian Random Wavefield	139
3.3.2. Statistics of the Complex Polarization Ratio for a Gaussian Random Wavefield	145
3.3.3. Statistics of Stokes Parameters for a Gaussian Field	147
3.3.3.1. Probability Density Functions of Stokes Parameters	147
3.3.3.2. Cumulants of Stokes Parameters	151
3.3.3.3. Probability Density Functions of Normalized Stokes Parameters	153
3.3.3.4. Statistics of Time-Integrated Stokes Parameters	158
3.3.4. Unpolarized Radiation: Stokes–Verdet–Barakat Conditions	160
3.3.5. Polychromatic Radiation Wavefield	164
<b>3.4. Entropy of the Radiation Field</b>	<b>165</b>
3.4.1. Entropy of a Stochastic Plane Wavefield	166
3.4.2. Temperature of Polarization	168
3.4.3. Maximum Entropy Principle	171
<b>Further Reading</b>	<b>176</b>
<b>PART 4 INTERACTION OF RADIATION WITH LINEAR MEDIA</b>	<b>179</b>
<b>4.1. Jones and Mueller Polarization Transfer Matrix Methods</b>	<b>181</b>
4.1.1. Jones Calculus	182
4.1.1.1. Eigenvalue–Eigenvector Decomposition	185
4.1.1.2. Jones Formalism and $2 \times 2$ Coherency Matrix Formulation	186
4.1.1.3. Gain	187
4.1.1.4. Jones Matrices for Common Optical Devices	191

4.1.1.5.	Physical Realizability Constraint	200
4.1.1.6.	Polar Decomposition of Jones Matrices	201
4.1.1.7.	Optical Propagation in Multilayered Media	203
4.1.1.8.	Polarization Ratio Transformations	206
4.1.1.9.	Jones Calculus and Entropy Transformations	208
4.1.2.	Mueller Calculus	210
4.1.2.1.	Mueller Formalism and $4 \times 4$ Coherency Matrix Formulation	211
4.1.2.2.	Gain	212
4.1.2.3.	Physical Realizability Conditions	213
4.1.2.4.	Mueller Matrices for Common Optical Devices	217
4.1.2.5.	Mueller Matrices and Their Parametrizations Within Poincaré Space	223
4.1.2.6.	Mueller Calculus and Entropy Transformations	227
4.1.3.	Relationship between Jones and Mueller Matrices	228
4.1.4.	Polarization Transfer in Nondepolarizing Optical Linear Media	233
4.1.5.	Mueller Matrix Analysis of Light Depolarization by a Linear Optical Medium	235
4.1.6.	Digression: Extended Matrix Formalisms	242
<b>4.2.</b>	<b>Polarization Effects at Dielectric Interfaces</b>	<b>244</b>
4.2.1.	Mueller Matrix Formulation of Fresnel's Equations	245
4.2.2.	The Pile-of-Plates Polarizer	247
4.2.3.	Principle of Ellipsometry	248
<b>4.3.</b>	<b>Polarized Light and Symmetry Transformations</b>	<b>252</b>
4.3.1.	Spatial Symmetry Relations for a Far-Field Scattering	253
4.3.2.	Time-Reversal Invariance and Reciprocity	260
4.3.2.1.	Time-Reversal Invariance	261
4.3.2.2.	Reciprocal Jones Medium Constraint	265
4.3.2.3.	Reciprocal Mueller Medium Constraint	267
<b>4.4.</b>	<b>Random Media</b>	<b>268</b>
4.4.1.	Jones and Mueller Matrices for Temporally Random Media	269
4.4.2.	Multiple Scattering by a Spatially Random Medium	275
4.4.2.1.	Rayleigh Scattering	277
4.4.2.2.	Monte Carlo Simulations of Wave Propagation Through Three-Dimensional Inhomogeneous Media	283

4.4.2.3. Backscattering Enhancement from a Random Distribution of Scatterers	293
<b>Further Reading</b>	<b>297</b>
 <b>PART 5 APPLICATIONS TO SELECTED TOPICS</b> <span style="float: right;"><b>303</b></span>	
<b>5.1. Electromagnetic Propagation in Linear Anisotropic Media</b>	<b>305</b>
5.1.1. Plane-Wave Propagation in a Linear Medium with Permittivity Tensor	306
5.1.2. The Differential Polarization Matrices	309
5.1.3. Ellipse of Polarization in Anisotropic Media	317
5.1.4. Evolution of Stokes Parameters in Anisotropic Media	320
5.1.5. Application to Liquid Crystals	324
<b>5.2. Optical Polarizing Components</b>	<b>330</b>
5.2.1. Polarizers	331
5.2.1.1. Reflection Polarizers	331
5.2.1.2. Transmission Polarizers	332
5.2.1.3. Polarization by Selective Absorption	334
5.2.2. Compensators	336
5.2.2.1. Fresnel Rhomb	336
5.2.2.2. Adjustable Compensators	338
<b>5.3. Measurement of Stokes Parameters</b>	<b>339</b>
5.3.1. Methods for Measuring Stokes Parameters	339
5.3.1.1. Stokes' Procedure	340
5.3.1.2. Collett's Procedure	341
5.3.1.3. Division-of-Amplitude Photopolarimetry	343
5.3.2. Probability Density Functions of Stokes Parameters	344
<b>5.4. Measurement of Jones and Mueller Polarization Matrices</b>	<b>346</b>
5.4.1. Methods of Determination of Jones Matrix	347
5.4.1.1. Jones Procedure	347
5.4.1.2. Differential Polarized Spectroscopy	348
5.4.2. Methods for Measuring Mueller Matrix	351
5.4.2.1. Transmission Experiment	352
5.4.2.2. Polarization Modulation Method	352
<b>Further Reading</b>	<b>357</b>

<b>APPENDIXES</b>	<b>361</b>
<b>Appendix A. Analogy of Second-Order Coherence Properties of Blackbody Radiation with Theory of Homogeneous and Isotropic Turbulence of an Incompressible Fluid</b>	<b>363</b>
<b>Appendix B. Derivation of Eq. (3.4.5) Using the Spectral Decomposition Theorem</b>	<b>366</b>
<b>Appendix C. Degree of Polarization of an Incoherent Mixture of Partially Polarized Light Beams</b>	<b>368</b>
<b>Appendix D. Set of Generalized Stokes Parameters</b>	<b>371</b>
<b>Appendix E. Short Historical Account of the Concept of Entropy of Radiation Fields</b>	<b>372</b>
<b>Appendix F. Singular Value Decomposition</b>	<b>374</b>
<b>Appendix G. Derivation of Bounds for Degree of Polarization</b>	<b>376</b>
<b>Appendix H. Application of Maximum Entropy Principle in Polarization Optics</b>	<b>378</b>
<b>Appendix I. Set of Coefficients <math>D_k^I</math></b>	<b>380</b>
<b>Appendix J. Evaluation of Correlation Function in Eq. (4.4.34)</b>	<b>382</b>
<b>Appendix K. Diffusing-Wave Spectroscopy</b>	<b>387</b>
<b>Appendix L. Polarization of Photons</b>	<b>392</b>
<b>AUTHOR INDEX</b>	<b>397</b>
<b>SUBJECT INDEX</b>	<b>401</b>

## PREFACE

---

Polarization of light is of considerable interest in our understanding of the physical world and has a broad range of technological applications. These include fields as diverse as optical communication, biological optics, atmosphere propagation, and display technologies. Although building on a foundation a century old, work in this field in the last decades has experienced substantial advances, including polarization aspects in multiple scattering by random media, and it is safe to project that this pace will be maintained far into the future. It seemed to me that no book that provides readers with a good understanding of polarization phenomena described by statistical optics has been published yet. I have tried to fill this gap.

It is commonly believed that one begins to understand a subject thoroughly only when teaching it to others. This book, which evolved from a set of course notes developed at Fourier University in Grenoble and at the University of Brest, France, is addressed principally to researchers already engaged in or wishing to enter the field of polarization optics as well as to graduate students in physics and optics. I hope that this book will also benefit those segments of the industrial and academic communities whose perception of polarization optics may not have evolved at the same pace as this discipline. Central ideas and theoretical approaches, rather than details of the experimental applications, are emphasized. One purpose of this book is to present some of the earlier developments. To avoid duplication, no attempt was made to cover ellipsometry (see, e.g., Azzam and Bashara [1]), photoelasticity (see, e.g., Theocaris and Gdoutos [2]), and crystal optics (see, e.g., Ramachandran and Ramaseshan [3] and Yariv and Yeh [4]). The present monograph inevitably builds on the achievements of past generations of workers in the field. Here we mention in particular the textbooks by Born and Wolf [5] and Shurcliff [6]. These sources were used in writing this book.

In this book, physical concepts and mechanisms are emphasized using appropriate mathematical details. My primary purpose is to provide a detailed analysis of those theoretical aspects of polarization optics that are capable of experimental verifications. Certain mathematical details are essential here because they are the most effective language for expressing the theory. Faraday in a letter to Maxwell [7] gave a clear statement as to the meaning of the

mathematical rigor we used. We can do no better than to quote him on this point:

When a mathematician engaged in investigating physical actions and results has arrived at his own conclusions, may they not be expressed in common language as fully, clearly, and definitely as in mathematical formulae? If so, would it not be a great boon to such as we to express them so—translating them out of their hieroglyphics that we also might work upon them by experiment. I think it must be so, because I have always found that you could convey to me a perfectly clear idea of your conclusions, which, though they may give me no full understanding of the steps of your process, gave me the results neither above nor below the truth, and so clear the character that I can think and work from them. If this be possible, would it not be a good thing if mathematicians, writing on these subjects, were to give us their results in this popular useful working state as well as in that which is their own and proper to them?

It was with this thought in mind, that this book was conceived and developed. Ultimately, the reader will decide whether I have succeeded. I assume only some familiarity with the basic notions of matrix algebra and elementary acquaintance with the theory of random processes, but no more than is absolutely necessary. The theory discussed in this book is exposed in classic terms almost exclusively. For a detailed exposition of the quantum description of the electromagnetic field, the reader should consult the textbook by Louisell [8].

The basic plan of the book is as follows. The book is divided into five parts. In Part 1, we go a little over the history of the subject and review the major milestones of polarization optics over the decades, with special attention to the individuals responsible for particular discoveries. Concepts rarely arise out of nowhere. There is always an historical context, in which isolated precursors of an idea have already appeared. On one hand, I believe that exploring the historical growth of ideas from which our contemporary views of polarization were fashioned forms an important part of human understanding of modern physical sciences. On the other hand, there is no substitute for reading the original papers. To quote Maxwell [9] “it is of great advantage to the student of any subject to read in the original...for science is always most completely assimilated when it is found in its nascent state. Every student of science should, in fact, be an antiquary in his subject”. In this respect we have given a substantial bibliography. Part 2 discusses the classic radiation field theory in a standard presentation and serves as a brief introductory to Maxwell’s theory of electromagnetism. The formal description of partial polarization is then given in Part 3, which also covers the basic mathematical framework for describing the statistical properties of a radiation field. We also discuss the thermodynamic description of a radiation field and introduce the concept of polarization entropy. With the set of these theoretical tools, the effect of propagation of light through a linear optical system is discussed in general terms in Part 4. Polarization transfer matrix methods provide a full description of the changes

in polarization state that a beam of light undergoes in interacting with optical elements. This part includes a discussion of the polarization by reflection and transmission. Some technical points that are often taken for granted in the literature are elaborated, such as optical reversibility and symmetry relationships. A topic that logically fits in at this point is the study of the propagation of partially polarized waves in disordered systems. In particular, we discuss issues such as the evolution of a partially polarized wave in a temporally random medium and the effect of multiple scattering of light by a spatially random distribution of scatterers. Part 5 then treats several advanced topics for which there does not exist an extensive literature. The current spate of polarization-related papers is so broad that no one text can hope to cover all aspects and to describe their importance in other areas of physics, such as magneto optics. Like any discussion of a broad field, this one has its limitations. A process of selection was needed to keep the length of this book within reasonable bounds. Instead of discussing many topics briefly, I chose to discuss a few topics in some depth that should be useful to practitioners. The topics selected for discussion include description of the propagation of partially polarized light in anisotropic media and the measurement methods of the Stokes parameters and of the polarization matrices. These topics were chosen because we felt that readers could obtain a sufficient impression of the applications of these topics to appreciate the development of the theory. Many of the mathematical methods used in the main text are generally unfamiliar to optical physicists interested in the subject. For this reason there are also a series of appendixes to which readers may refer if necessary. Carefully selected references to the original literature for additional readings are given at the end of each part. The bibliography of articles and books has been assembled with the aim of giving accessible references that further illustrate the points made in the text. For this reason, I cited mostly research papers the purpose and approach of which are similar to those in that book.

This book is not an exhaustive survey of polarization optics, but rather its aim is to convey to the reader some of the latest understandings that pervade this exciting field of inquiry and that lead to more precise control of polarization in optical systems.

## REFERENCES

1. R. M. A. Azzam and N. B. Bashara, *Ellipsometry and Polarized Light*, North-Holland, Amsterdam, 1977.
2. P. S. Theocaris and E. E. Gdoutos, *Matrix Theory of Photoelasticity*, Springer-Verlag, Berlin, 1979.
3. G. N. Ramachadran and R. Ramaseshan, "Crystal Optics," in *Handbuch der Physik*, S. Flugge, ed., Springer-Verlag, Berlin, 1961, Vol. XXV/1.
4. A. Yariv and P. Yeh, *Optical Waves in Crystals*, Wiley, New York, 1984.
5. M. Born and E. Wolf, *Principles of Optics*, 6th ed., Pergamon Press, Oxford, 1980.

6. W. A. Schurcliff, *Polarized Light*, Harvard Univ. Press, Cambridge, MA, 1962.
7. M. Faraday, *Experimental Researches in Electricity: Faraday's Collected Works*, Dover, New York, 1964.
8. W. H. Louisell, *Quantum Statistical Properties of Radiation*, Wiley, New York, 1973.
9. J. C. Maxwell, *A Treatise on Electricity and Magnetism*, 2 vols, Dover, New York, 1962.

## ACKNOWLEDGMENTS

I should like to thank the many people that made this work over the years possible. Discussions with Dick Barakat at Harvard University, Ed Rockower at the Naval Postgraduate School, Monterey; Dominique Bicout at the Institut de Biologie Structurale, Grenoble; and Alex Kostinski at the Michigan Technological University were most useful at several steps of this work. The work of other individuals is gratefully noted. I am indebted to Professor R. A. Chipman at the University of Alabama in Huntsville for reading parts of the manuscript and making many suggestions. My thanks go to Professor Craig F. Bohren at the Pennsylvania State University for his critical comments and numerous helpful suggestions. In the planning and execution of this project, I enjoyed the understanding and the patience of Greg Franklin, John P. Falcone, and Rosalyn Farkas, at John Wiley. Needless to say, none of these people should be blamed for the shortcomings of the final product.

CHRISTIAN BROSSEAU

Brest, France  
June 1998

## PART 1

---

# HISTORICAL SURVEY OF UNDERSTANDING OF POLARIZED LIGHT

For centuries, the peculiar polarization properties of light have proved fascinating. To make the connection with the way we understand things nowadays and before trying to explore such questions as what polarized light is and how it works, I believe it is first necessary to recount the major contributions to the subject. Newcomers to the subject cannot understand the significance and relevance of a discovery if they do not have the proper historical context in which to place it.

The study of polarized light and its interaction with matter has been an area of continuing interest for physicists, and its understanding has challenged the human mind through the ages [1–9]. The origins of the modern concept of polarization can be found in the scientific literature of the late nineteenth and early twentieth centuries. Some early stages in the development of understanding polarization optics from the *experimentum crucis* of 1669 on are recalled in this narrative. In this inventory we cannot look at the full history and extent of research using polarized light. Rather, we have chosen to confine ourselves to the presentation of the chief historical events over a period of three centuries with an assessment of how each contribution helped to develop the overall understanding of polarization. We are interested in the scientific concepts and how they were ultimately united into a coherent theory. Some of the important issues we address are when it started and who provided the knowledge base? The organization of the experimental and theoretical material tries to follow chronological order. We are forced to recognize that the progression in the process of gaining knowledge presented here is an order imposed in retrospect by the author. We have chosen arbitrarily to close this brief survey at the beginning of the development of quantum and nonlinear optics, which are outside the scope of this book. We will see that the list of contributors includes some of the greatest names in physics. In most of these cases the motivation of the research work was fundamental knowledge.

The survey undertaken here discusses in the first section the origins of the basic concepts. The second part focuses on the discovery of the electromagnetic nature of light, while the last section deals with the understanding of coherence

## **2 HISTORICAL SURVEY OF UNDERSTANDING OF POLARIZED LIGHT**

and quantum properties of light. Dates refer to the fundamental papers or reports. Ample references to most original literature are given at the end of this chapter for those who desire source material and those who want to place in context and appreciate the considerable achievements of people who have contributed solidly in their own fields. Little detail is provided on personal lives; rather emphasis is placed on the technical developments. It is an entirely subjective account since it is a difficult task of appropriately weighting individual contributions in retrospect.

## SECTION 1.1

---

# First Period: Early Ideas and Observations—from Bartholinus to Stokes

Let us begin our excursion by reviewing a few facts about the scientific context of the seventeenth century. At that time optics was crucial for the intellectual development of Europe, and purely geometric (Euclidian) considerations sufficed to explain most major optical observations, specifically, the rectilinear propagation of light through transparent media (Fermat's principle). Much less had been formulated concerning the physical principles underlying the very nature of light. Light was considered as the instantaneous propagation of an action (of mechanical origin) through a kind of subtle ethereal matter (luminiferous aether) that fills all space [4]. It is in this context that the story of polarization optics starts. It is rather difficult to tell who was the first in polarization optics. Although it is a part of the established folklore that the Vikings (circa 700) might have used polarization of skylight to find their way across the Atlantic using the famed sunstone mentioned in the medieval sagas [10], this belief should be treated with extreme caution [11]. The discovery of polarization is usually attributed to Bartholinus [12]. Erasmus Bartholinus, a Danish mathematician at the University of Copenhagen, is credited with the first, scientifically presented, observation of a polarization effect. In 1669 he made the pivotal discovery of polarization by double refraction using a crystal of Iceland spar (i.e., calcite, rhombohedral crystalline form of calcium carbonate  $\text{CaCO}_3$ , was discovered in Iceland and was then known as *Iceland spar*) and published his observations in a 60-page memoir entitled "Experimenta crystalli Islandici disdiaclastici quibus mira et insolita refractio delegitur" [12]. In fact, he observed the splitting of an incident beam into two beams of equal intensity after passing through a crystal. One of these rays—the *ordinary ray*—obeys the usual law of refraction; the other does not and hence is called the *extraordinary ray*. At this point, it is worth noting that the challenge of accounting for the optical properties of crystals gave the starting impetus to the study of polarization optics.

In 1672 Christiaan Huyghens, a Dutch physicist, interpreted the double-refraction phenomenon from his conception of a spherical lightwave (i.e., envelope construction) and observed that each of the two beams arising from

the double-refraction phenomenon can be extinguished by passing through a second calcite crystal that is rotated about the direction of the beam [13]. His investigations also showed that the two beams have different polarization directions and that to prevent the separation of light, light must traverse the crystal parallel to the crystallographic axis. He subsequently completed in 1690 his famous book on light, *Traité de la Lumière*, in which he developed the geometric theory required to interpret all optical phenomena known to this time: reflection, refraction, and double refraction [13]. His reasoning was based on the principle named after him, according to which light behaves like a spherical wave emitted from a source; each point of the ether on which light falls becomes the source of spherical wavelets (secondary waves) that propagate in all directions and that combine to form the wavefront (the envelopes of the spherical waves) perpendicular to the direction of propagation. That was a major conceptual advance. Note that the finiteness of the velocity of light demonstrated by Ole Römer in 1676 was one of the two basic ingredients of Huyghens' analysis; the other was the wave nature of light [3, 4].

In 1704 Sir Isaac Newton published his celebrated treatise on optics (*Opticks*), which was the standard treatise on this subject for many decades [14]. Newton was one of the most famous “emisionnists”—a supporter of the corpuscular or particle theory of light [5]. In this view a beam of light comprises many rays, identified with geometric lines, which are, in turn, composed of streams of particles. The differences in colors come from the differences in particle size and mass. In this scheme refraction and reflection are understood in terms of Newtonian forces acting between these particles and the interface. Since these ideas were embedded in his theory of gravitation (a theory that was verified in 1758 by the return of the Halley comet, predicted in 1705), it was essentially this view that prevailed through the eighteenth century because of the great prestige lent it by Newton.

The seed of a scientific breakthrough may slumber for a time without generating much interest. It is not until the start of the nineteenth century that we see the next progress in the field. This period marked the turning of the tide in favor of the view that light vibrations are transverse to the direction of propagation, the evidence for which had been advanced vigorously by Thomas Young [15]. The double-slit optical interference experiment of Young performed in 1801 was a key discovery to the understanding of the wave theory of light, according to which he established that beams of light can constructively and destructively interfere. Young's experiments were crucial because destructive interference was not a possibility according to the Newtonian corpuscular theory. Another big advance in the understanding of the nature of light came when Thomas Young showed in 1803 that polarization phenomena arise from the transverse nature of light. The latter recognition gave impetus to the wave theory of light. Bear in mind that this theory predates the electromagnetic theory of light by more than six decades.

The next great progress came with Malus. In 1808 Etienne-Louis Malus, a military French engineer, discovered the polarization of natural light by

reflection while experimenting with a crystal of Iceland spar and light reflected by the windows of the Palais du Luxembourg in Paris. By extensive experimentation, he showed by purely geometric reasoning how to express the intensity of light emerging from a polarizing crystal when the light it receives is linearly polarized along a direction making a specific angle with its axis, in a paper entitled “Sur une propriété de la lumière réfléchie” [16]. It has been common to give to physical effects names that honor the scientists who first draw attention to them. This discovery is no exception and is known as *Malus' law*.

In 1812 Sir David Brewster, a Scottish clergyman-turned-physicist at Saint Andrews College, enunciated the tangent law named after him in a couple of celebrated papers: “On the laws of polarization and double refraction in regularly crystallized bodies” and “On the laws which regulate the polarization of light by reflection from transparent bodies” [17]. What Brewster did was to show how the polarizing angle (the angle of incidence at which unpolarized light becomes completely polarized upon reflection by an optically smooth, planar interface between two dissimilar transparent media) depends on composition (by way of refractive index). Like Malus, he investigated polarization by reflection of natural light by planar interfaces between two media. One of his early achievements was his discovery that at a particular angle of incidence, the reflected light from glass after passing a calcite crystal could be extinguished. This angle is called the *Brewster angle*. Moreover complete polarization occurs when the angle of incidence equals Brewster's angle. Photoelasticity (strain birefringence) was also described by Brewster in 1816: the possibility that isotropic transparent optical media could be made optically anisotropic by the application of mechanical stress is a topic of profound significance on which Brewster has had a seminal influence. The theory was worked out subsequently by Neumann [18]. Brewster also invented and patented the kaleidoscope [17].

A major advance in the understanding of light polarization was made by Augustin Jean Fresnel. In 1823 he derived, on the basis of the elastic theory of ether (the concept of an ethereal medium, filling space, was formulated by Descartes two centuries before), his famous reflection and transmission formulas for a plane wave that is incident on a static and plane interface between two dielectric isotropic media in an article entitled “Mémoire sur la double réfraction que les rayons lumineux éprouvent en traversant les aiguilles de cristal roche suivant des directions parallèles à l'axe” [19]. It is important to note that Fresnel's equations require arbitrary sign conventions that may lead to confusion: a fact already mentioned by Fresnel as the “petite difficulté.” Fresnel gave also a heuristic explanation of the phenomenon of optical activity and recognized the phenomenon of optical rotation arising from circular birefringence, namely, that such material has an index of refraction for right circularly polarized light which is different from the index for left circularly polarized light. His pioneering work concerning the diffraction caused by small apertures and screens is also fundamental. To Fresnel must be assigned credit

for discovering the modern concept of polarization of light and stimulating the efforts which put the wave theory of light on a firm foundation.

In the years 1812–1815 came the important milestone by the French physicist Dominique François Arago at the Paris Observatory. Arago performed many fundamental investigations on optical activity in quartz: the continuous rotation of the direction of vibration of light by propagation along the optic axis of a crystal. Arago invented the pile-of-plates configuration, sometimes credited, incorrectly, to Stokes, whose paper on the pile-of-plates polarizer followed Arago's by more than 20 years [20]. On the theoretical side, his principal contribution was discovery of the interference laws published in a joint paper with Fresnel, "Sur l'action que les rayons lumineux polarisés exercent les uns sur les autres," which played a key role in the demonstration of the transverse nature of lightwaves propagating in free space [21]. Arago and Fresnel's contribution was to enunciate four interference laws governing the interference of polarized light. These laws were determined experimentally and are summarized by the following statements (in modern version): (1) two waves, linearly polarized in the same direction, can interfere; (2) two waves, linearly polarized with orthogonal polarizations, cannot interfere; (3) two waves, linearly polarized with orthogonal polarizations, if derived from orthogonal components of unpolarized light and then brought into the same plane of oscillation, cannot interfere; and (4) two waves, linearly polarized with orthogonal polarizations, if derived from the same linearly polarized wave and then brought into the same plane, can interfere [22].

In 1815 there soon followed the discovery by Jean Baptiste Biot of the optical activity of some materials, described in the article "Sur un mode particulier de polarisation qui s'observe dans la tourmaline," and performed experiments of rotatory polarization to distinguish between left- and right-handed rotation of the direction of vibration of light [4]. Shortly after, in 1822 Sir John Herschel, astronomer at Cambridge University, presented evidence that the rotatory polarization of quartz originates from two different crystallographic structures: a phenomenon called *enantiomorphism* [23].

In point of fact the first decades of the nineteenth century were a very exciting time in optics. Many of the basic designs of innovative optical instruments were improved and made many discoveries possible. For example, the development of the optical microscope opened new realms to human investigation. This encouraged opticians to make numerous measurements on a wide variety of systems and at the same time made it possible to observe many subtle effects. Many landmarks ideas come out of these studies. The first birefringent polarizer was realized in 1828 by the Scottish physicist William Nicol, at the University of Edinburgh. Nicol is known as being the inventor of the polarizing prism that now carries his name, described in an article "On a method of so far increasing the divergency of the two rays in calcareous-spar that only one image may be seen at a time" [24]. The first polarizing microscope was built by H. F. Talbot in 1834 and used by Brewster for studying minerals. Several ingenious optical designs have been carried out in

an attempt to improve the quality of measurements. Significant contributions were made, including the polariscope in 1830 by A. Seebeck, the “pince à tourmalines” by J. Müller, the compensators by Babinet and Senarmont, the Nörrenberg reflecting polariscope and the Glan-Foucault polarizers, and the Wollaston and Rochon prisms [4]. In 1846, Wilhelm Haidinger, an Austrian mineralogist, discovered the remarkable brush phenomenon that now bears his name. This optical curiosity is observable and apparent to almost anyone: it is manifested as specific patterns that can be seen when one looks at polarized skylight [25, 26]. The following year, Haidinger was the first to report an observation of circular dichroism in an article “Über den pleochroismus des amethysts.”

The first to realize that molecular symmetry is responsible for the phenomenon of optical rotation is Louis Pasteur in his doctoral research done in 1848 [27]. Examining different salts of tartaric acid, he noticed the formation of two types of crystals, each one a mirror image of the other (enantiomorphs). Separating the two types of crystals and then dissolving these separately in water, he found that one solution caused the direction of polarization of linearly polarized light to rotate clockwise and the other, counterclockwise. The remarkable discovery of the ability of left-handed and right-handed molecules to rotate light differently is at the basis of what is now termed *chiral asymmetry*, which plays a fundamental role in the chemistry of life. It is likewise instructive to note that William Bird Herapath, physician and surgeon in Bristol, England, observed for the first time linear dichroism using crystals of quinine sulfate periodide (now known as *herapathite*). The high point of Herapath’s scientific career was the discovery in 1852, published in the *Philosophical Magazine* under the title “On the optical properties of a newly-discovered salt of quinine which crystalline substance possesses the power of polarizing a ray of light, like tourmaline and at certain angles of rotation of depolarizing it like solenite” [28].

A major advance to the field came by Sir Georges Gabriel Stokes. Stokes introduced four measurable quantities that now bear the name of Stokes parameters for describing the properties of polarized light. He published the seminal paper, “On the composition and resolution of streams of polarized light from different sources”, in 1852 [29]. Despite the impressive insight and relatively simple concepts underlying Stokes’ original paper, this remarkable study received little attention at that time but would become fully appreciated many decades later. Although his basic motivation for the introduction of these four parameters was to describe mathematically unpolarized light, his parametrization is applicable to any state of polarization; to partially polarized, as well as unpolarized and completely polarized, waves. In a sense the modern description of polarization optics can be said to have started from the time of the publication of papers by Stokes. This was an entirely new and unexplored point of view because his description of light was formulated in terms of intensities rather than field vectors, thus dealing with measurable quantities at optical frequencies. The first parameter represents the total intensity of the

field, and the three others describe the state of polarization. He also formulated a mathematical statement for unpolarized light: the intensity of unpolarized light is unaffected by any rotation of the axes of the reference coordinate system and to any phase change introduced in one of the wave components. Stokes was eventually the best interpreter of the interference laws that Fresnel and Arago discovered. To polarization, Stokes brought also in 1849 the principle of reciprocity (reversibility) involving reflection and transmission of light beams on an interface between two dielectric media.

In 1860 Giorgio Govi, Italian astronomer, published his finding that the light scattered by small particles can be polarized (in his experiment, he filled a closed room with smoke and directed a beam of sunlight onto the smoke) and rediscovered several observations first made by Arago [30]. His conclusion, that scattering of light by small particles is a phenomenon that is independent of reflection of light (a common idea shared at the time), opened the science of the scattering of radiation by matter. The early work in environmental monitoring science was led by the Scottish physicist John Tyndall. His brilliant experiments demonstrating that the scattering of light by particles depends significantly on the size of the scatterers are particularly noteworthy. He was the first to find that a cloud of very fine particles perfectly polarizes the beam at 90° scattering angle. Tyndall's 1869 study of photochemical smogs published under the title "On the blue color of the sky, the polarization of sky light and on the polarization of light by cloudy matter in general" would be quite decisive to the subsequent analysis of Rayleigh [31]. More specifically, he was the first to observe the bluish color of light after scattering of natural light. Moreover, he invented many instruments and used them to conduct nephelometric research into the color of water and sky.

Let me interrupt the story here to say something about the subject before the entry of the seminal electromagnetic theory of Maxwell. Two features are of special pertinence to the following discussion. First, the transverse wave theory of light could provide an understanding of the major optical phenomena discovered at the time of Maxwell's treatise: propagation, polarization, diffraction, and interference. The majority of scientists became convinced of the superiority of the wave theory in the late 1820s and early 1830s even though this theory said nothing about the source of the optical field [5]. Second, in spite of many difficulties, the mechanical theory of the elastic ether persisted.

---

## SECTION 1.2

---

# Second Period: The Electromagnetic Nature of Light

A new chapter in this story was the entry of Maxwell's theory of electrodynamics. The electromagnetic nature of light was established, and the connections between radiation of many different kinds was discovered. The story involved many of the intellectual giants of the nineteenth century. This period began with the Danish physicist Oersted in 1820. By experimenting the effect of a current on a suspended magnetic needle, he made the fundamental discovery that magnets interact with electric currents [4].

The second third of the nineteenth century brought a profound change in physics. A chief contributor to the fundamental aspects of early electromagnetic theory was the British physicist Michael Faraday. Faraday was one of the greatest experimental scientists of all time. In 1831, he discovered electromagnetic induction, subsequently explained para- and diamagnetism and interpreted them through his field theory. When Faraday postulated the physical laws of electromagnetism, he had in mind a mechanical picture, deduced from geometric reasoning, using the two concepts of lines of force traversing all space and actions-at-a-distance exerted between the particles in a medium [32]. In 1845 he discovered the phenomenon, which now bears his name, of magnetically induced optical rotation of the direction of vibration of light on passing through a medium in a direction parallel to the magnetization. His eponymous law has become the earliest indication of the relationship between electromagnetism and light [32]. According to Einstein, Faraday (along with Maxwell) was responsible for the greatest change in the axiomatic basis of physics since Newton [33].

In 1864, James Clerk Maxwell completed a six-page memoir entitled "A dynamical theory of the electromagnetic field," in which he developed the mathematical theory required for the description of how electromagnetic waves propagate [34]. The formulation of Maxwell's theory certainly ranks as one of the prime events of nineteenth-century physics; his eponymous equations have become almost as immutable as Newton's law of gravity. These equations summarize the fundamental relations between electricity and magnetism and became the cornerstone on which generations of scientists have based their theoretical studies. Maxwell put Faraday's concepts into the elegant mathematical form of four differential equations, and one of his major innovations

was to introduce the notion of displacement current. Indeed, were it not for the displacement current, it would not be possible to deduce from Maxwell's equations that electromagnetic waves have the property of light. In the preface to his monumental *Treatise on Electricity and Magnetism*, he observed that "As I proceeded with the study of Faraday, I perceived that his method of conceiving the phenomena was also a mathematical one, though not exhibited in the conventional form of symbols." By Maxwell's time, scientists distributed the results of their research by means of correspondence. The reader may wish to consult Harman for an interesting discussion of letters Maxwell wrote to Faraday where he described his incorporation of his mechanical concepts to the theory of electromagnetic field [35]. According to Maxwell's mechanical model, the electric field represents a physical stress in the ether and the magnetic field represents the rate in change in the stress field. Thus electromagnetic radiation is a result of electric oscillations. The Maxwell's equations allow for the possibility of transverse wave solutions, and the corresponding phase velocity, expressed in terms of electric and magnetic quantities, is remarkably close to the free-space speed of light. Moreover, the transversality of the vibrations is consistent with the observation of the Arago–Fresnel interference experiment. Electromagnetism is often viewed as a concept-driven revolution in the sense proclaimed by Thomas Kuhn, that is, a revolutionary new way of looking at nature [36]. The determination of Rudolph Kohlrausch and Wilhelm Weber of the velocity of the electromagnetic waves was found in accordance with the velocity of light carried out by Henry Fizeau. The principle then emerged that light has the properties of electromagnetic waves [4, 37]. These fundamental researches constituted the seed corn that has led to a scientific harvest of inestimable value today.

With the basic physics introduced by Faraday and Maxwell, the theory of electrodynamics followed the development of the necessary mathematics with contributions from Poisson, Green, Thomson (Lord Kelvin), and others [4, 8]. Although now largely forgotten, significant progress in the study of polarized light was reported by Emile Verdet in 1869 [38]. Verdet may be singled out for mention as having first developed a rigorous analysis, surprisingly modern in tone, of unpolarized light. Another accomplishment of Verdet in this area concerns the Faraday effect. Verdet's constant of proportionality links the angle of rotation and the product of the magnetic field strength with the distance that the light travels through the medium. Also noteworthy are the investigations of unpolarized light by J. Stefan published in 1864 [39].

On the experimental side, a large number of important experiments were conducted, permitting the observation and the study of natural optical rotation, natural and magnetic linear, and circular dichroism. In 1875, John Kerr, Scottish physicist, discovered the quadratic electrooptic effect that bears now his name. Indeed, the full title of Kerr's article was: "A new relation between electricity and light: dielectrified media birefringent" [40]. By experimenting with a block of glass between crossed polarizers and applying an electric field

( $\sim 10^5 \text{ V m}^{-1}$ ) normal to the optic axis, Kerr found that the glass becomes birefringent (proportional to the square of the field) and part of the light is transmitted. In 1893 the German physicist Friedrich Karl Pockels discovered a new (linear) electrooptic effect—in the Pockels effect, the field is parallel to the direction of light. When a dc (direct-current) electric field ( $\sim 10^5 \text{ V m}^{-1}$ ) is applied to a noncentrosymmetric crystal, such as lithium niobate ( $\text{LiNbO}_3$ ), the refractive index changes linearly with the applied voltage and incident linearly polarized light becomes elliptically polarized [40]. A variety of other magnetooptic phenomena were discovered over the years. The Kerr magneto-optic effect is the analog of the Faraday effect in the case of light reflected off a magnetic or magnetized material; the Cotton-Mouton effect is double refraction of light passing perpendicular to the magnetic field applied to the material, that is, the magnetic equivalent of the Kerr effect [40]. Materials that show large electrooptic and magnetooptic effects have become important in the transmission, storage, and retrieval of digital data. A major experiment was performed by O. Wiener in 1892 that demonstrated that the electric field vector of an electromagnetic wave exerts a greater force on electrons than the magnetic field does (provided one considers nonrelativistic speeds). Thus the electric field is responsible for changes of polarization state on the interaction of light with matter. This discovery is at the basis of the convention to chose the direction of the electric field as the polarization direction of the lightwave [41].

A great experimental support to Maxwell's theory of electromagnetism came with Hertz' experimental work. In the years 1887 and 1888 Heinrich Rudolf Hertz, a German physicist at the Technical University in Karlsruhe, produced and detected electric waves in air, thus demonstrating the application of the concepts of the electromagnetism theory to the microwave and radio regions of the spectrum. Among Hertz' achievements was the experimental confirmation of the transverse nature of electromagnetic waves through experiments on reflection, refraction, interference, and polarization, in accordance with the predictions of Maxwell's theory [42, 43]. Hertz' experimental work lead to the technology of radio and microwave communications, detection, and wire-grid polarizers to test the properties of radiowaves. In addition to being a remarkable experimentalist, he was a first-rate theoretical physicist; his fundamental work helped lay the foundation for quantum theory and relativity. On one hand, Hertz' observation of the photoelectric effect in 1887 played an important role in the development of quantum physics, and his work on cathode rays with Lenard was crucial for Wilhelm Roentgen's discovery of X rays in 1895. On the other hand, Hertz' investigations into the electrodynamics of moving bodies advanced ideas that eventually led to Einstein's special theory of relativity. In the late 1890s appreciation of the significance of the electromagnetic theory was increasing. Within this context, in 1896, Zeeman identified a fundamental rule concerning the broadening of the yellow D lines of sodium atoms in a magnetic field. Then, on the basis of the experiments by Zeeman, Hendrik Antoon Lorentz quickly developed a theory and was able to predict

the polarization behavior of the spectral lines. These predictions were experimentally verified by Zeeman. Thus, all optical phenomena can be accounted for by the interaction with electromagnetic fields. This success had a vast and immediate impact on the acceptance of the electromagnetic theory of light within the optics community. Indeed, Maxwell's successful theoretical explanation of the observed phenomena of light propagation did not lead to a universal acceptance of the electromagnetic hypothesis, which was by no means evident to many leading scientists in the end of the nineteenth century [5]. That work won Lorentz and Zeeman the 1902 Nobel prize in physics.

Several other works are noteworthy. The reciprocity theorem, due to Hermann von Helmholtz in 1881, is fundamental in scattering processes. It expresses the relationship existing when the incident and scattered beams are interchanged. Around 1889 Drude rederived Fresnel's formulas from Maxwell's equations and this constitutes another important step in the acceptance of Maxwell's theory [4–6]. A major breakthrough was achieved in 1890 with the publication of a 96-page memoir, "Upon the reflection and refraction of light by a nonmagnetic transparent homogeneous sphere," of arbitrary size and refractive index by the mathematical physicist Ludwig V. Lorenz of Copenhagen, which predated the classic work of Mie (1908) about particle light scattering and absorption [44].

History justly positions the French mathematician and professor of mathematical physics at the Sorbonne, Paris—Henri Poincaré—as a dominant figure in mathematical physics. His ideas, discoveries, and techniques pervade all areas of theoretical physics. In 1892 Poincaré published a mathematical treatment of polarized light where he introduced the Poincaré sphere and the complex plane representations to specify the state of polarization (a convenient summary of much of his work is contained in his magnum opus on optics, *Théorie Mathématique de la Lumière*, which was the standard treatise on optics for many years) [45]. It was Poincaré, more than any other, who truly saw the physical implications of geometry in polarization optics. Using a stereographic projection, he mapped each point on the plane into a sphere whose points are in one to one correspondence with all the possible states of polarization of a light beam. One of the conveniences of the Poincaré sphere is that it provides an intuitively geometric view of the transformation of a polarized light when it interacts with optical devices in terms of rotations of states. Oddly enough, the importance of the Poincaré sphere in polarization optics was not realized until the second half of the twentieth century. Poincaré's large body of work on optics is a prime example of his elaboration of ideas in new fields. It is remarkable how many important ideas he put forward, establishing the basic concepts of modern chaos and dynamical systems theory. He was a prolific author (Poincaré's long and diverse bibliography contains more than 500 articles) and wrote several fine books on the theory of differential equations and dynamics, which were widely used texts. His accomplishments as a theoretical physicist make best sense when they are placed in this broader perspective of his scientific interests, giving any modern scientist ample cause for inspiration.

At the turn of the (nineteenth–twentieth) century, John William Strutt, the third Baron Rayleigh, more familiar as Lord Rayleigh, Cavendish professor of experimental physics at the Cavendish Laboratory in Cambridge and then professor of natural philosophy at the Royal Institution of Great Britain in London, published many fascinating articles in optics. Although he received the Nobel prize for physics in 1904 for his discovery of argon, Rayleigh had a long and extraordinary productive career. Lord Rayleigh contributed to almost every branch of classic physics: optics, acoustics, electromagnetism, thermodynamics, statistical mechanics, and others. His collected papers contain more than 400 papers. An important characteristic of Lord Rayleigh's way of doing physics was his close attention to experiment [46]. One of his major contributions came in 1871 in his famous paper entitled "On the scattering of light from small particles," when he derived the polarization at  $90^\circ$  law (polarization of the light scattered by the atmosphere when viewed at a right angle to the direction of incidence), the inverse fourth-power law for the intensity of light scattered by particles, not necessarily spherical, whose size is much smaller than the wavelength of the light and explained that the degree of polarization of the scattered light depends on the angle of scattering from the elastic-solid theory of the luminiferous ether [46]. Single scattering from air molecules is known to result in the polarization of scattered light in a cloudless blue sky.<sup>1</sup> In 1881 he rederived these results (for dielectric cylinders and small spheres) from the Maxwell's theory of electromagnetism, which, in fact, did not change his major conclusions. For a more detailed exposition of Rayleigh's legacy in physics, the reader should consult Ref. 46. In his book on the Theory of Sound, he established a principle of reciprocity in scattering which was subsequently extended to optics by Perrin.

Several major advances, both theoretical and experimental, quickly followed Rayleigh's work. Love investigated the problem of scattering of electromagnetic waves by a sphere. Then Gustav Mie in 1908 ("Beiträge zur Optik trüber Medien, speziell kolloidaler Metallösungen") and independently P. Debye ("Der Lichtdruck auf Kugeln von beliebigem material") in 1909 found an analytic solution to the scattering problem of a plane, time-harmonic electromagnetic wave of arbitrary polarization and frequency by a homogeneous sphere of arbitrary size provided that the sphere material is isotropic and is characterized by a permittivity that can be complex and frequency dependent, usually referred to as Mie theory [47, 48].

A list of classic highlights in polarization optics would not be complete without tribute to the following research contributions. The effect of magnetically induced birefringence in colloidal suspensions of ferromagnetic particles was first investigated by Majorana in 1902 [49]. In 1905 Umov reported on polarization aspects in the reflection of light by rough surfaces. Since their initial discovery in 1888 by Reinitzer, liquid crystalline materials have been

<sup>1</sup>Professor Craig Bohren kindly informed readers that the term *Rayleigh scattering* may be a source of confusion because of the varied range of phenomena with which it can be associated. The interested reader may consult C. F. Bohren, *Opt. Phot. News* **3**, 1992, 18 and A. Young, *Appl. Opt.* **20**, 1981, 533.

investigated intensely for their scientific and technological potentials. Perhaps the earliest instance of the important polarization aspects in liquid crystals was presented by Mauguin in his article entitled “Sur les cristaux liquides de Lehmann” [50]. He is given due credit as the first to present a Poincaré sphere construction (which was new at the time) to study the deformation of twisted nematic liquid crystals. Schuster and Schwarzschild initiated the first radiative transfer theory in 1905 for explaining the appearance of absorption and emission lines in stellar spectra [51]. In 1908 George Hale reported early interesting observations on the polarization of sunlight and detected circular polarization in the wings of a Fraunhofer line from a sunspot [51]. Significant improvements in optical instrumentation were made in the first decades of this century, including the polarizing microscope. Circular dichroism in solutions was extensively investigated by Aimé Cotton and Georges Bruhat [53]. As an aside, the finding of the circular polarization properties of certain families of beetles by A. Michelson in 1911 is noteworthy. He observed colorful effects resulting from the property of converting unpolarized incident light from the sun into left-circularly polarized reflected light [54].

During the end of the nineteenth century there was excitement about investigations of different kinds of radiation. In 1895 Wilhelm Conrad Roentgen, a professor of physics at the Julius Maximilian University of Würzburg, Germany, published his famous article entitled “Eine neue Art von Strahlen” in which he presented the discovery of high-energy electromagnetic radiations: the X rays (X for unknown). Note that the existence of “invisible” high-frequency electromagnetic radiations had been predicted by von Helmholtz a few years earlier [55]. That work won Roentgen the 1901 (and first) Nobel prize in physics. From Roentgen’s discovery to the present time, X-ray crystallography proved essential in relating the atomic structure of condensed matter to their functions and physical properties. The transverse nature of X rays was first described and demonstrated in a paper “Polarization in secondary Roentgen radiation” published in 1906 by the Scottish physicist Charles G. Barkla [56]. This advance was recognized (11 years later) when Barkla was the recipient of the 1917 Nobel prize for physics.

Before continuing with this narrative, it is perhaps appropriate to recall a few important facts. By the end of the nineteenth century and the beginning of the twentieth century, we are in a time of great change. Maxwell constructed a theory of electromagnetic radiation which was successful in explaining all the observed phenomena of propagation and scattering of light, but the detailed understanding of the processes of emission and absorption by matter required one more step in the theoretical description of the interaction between field and matter. An essential question that arose was this: What is the origin of light? The question can be clarified only by an appropriate quantum treatment.

## SECTION 1.3

---

# Third Period: The Coherence and Quantum Properties of Light

The development of quantum theory in the first third of the twentieth century brought a profound change in physics and had a major impact on our notions of the nature of light. The theory of electromagnetic radiation in thermal equilibrium was a source of mystery in the context of the physics prevailing at the turn of the century. In addition to providing clues about the physical principles underlying all optical phenomena, the depth of the new ideas (dual nature of light as both corpuscle and wave) provided by the quantum theory in 1924–1925 soon followed by quantum field theory in the late 1920s, sparked a revolutionary way of looking at nature, a paradigm change.

In the years up to about 1905 the subject of polarization optics seemed to have reached a sort of maturity; the impressive success of Maxwell's electrodynamics made it clear to optical physicists that the laws behind the phenomena had been comprehended. At the same time researchers developed a variety of clever methods for the analysis of polarization of light. A large fraction of the efforts in the field were being devoted to the study of novel techniques and instruments rather than to elucidation of the theoretical foundations concerning the nature of light. In spite of these efforts, it appears that the need to introduce a mathematical formalism for describing radiation fluctuations, particularly the correlation between the electric field at two space points, was fundamental to an understanding of the nature of light. This subject was taken up by several workers. Max Planck, one of the towering figures who built the edifice of quantum mechanics, gave a precise expression of the entropy of a polarized “pencil of radiation” [57], although this concept was first conceived by Wien in 1894. The statistical description of optical fields played a key role in the development of quantum physics and had a central role in the efforts to understand the nature of light. The origin of the subject lies in von Laue’s (who was Planck’s student) pioneering work in 1907 [58]. Prior to this time there had been no clear recognition of the fact that the variation of the electric field vector of light is a random process in both time and space. In a paper devoted mainly to the thermodynamics of radiation, Max von Laue introduced a measure of the coherence between two pencils of radiation. Although von Laue’s ideas diffused rapidly, they apparently did not receive the amount of attention they deserve since the next step in coherence theory was done two decades later.

By the early 1920s much had changed. Many scientists had become seriously concerned with the analysis of radiation fluctuations. This shift of understanding and attention may eventually be regarded as one of our century's pivotal developments in polarization and coherence theories. Developments in statistical optics were then proceeding at an impressive rate. In 1926, Max Berek introduced another measure of the coherence of light, which was called the *degree of consonance* [59]. Another highly influential result was the introduction of the density matrix formalism by John von Neumann, which has much to do with the coherency matrix formalism pioneered by Wiener [60]. Another notable contribution in coherence theory was the work of Lakeman and Groosmuller [61]. Later van Cittert calculated the second-order correlation at points on a screen illuminated by an incoherent quasimonochromatic source [62]. A great step was taken in 1938 when Frederik Zernike, a Dutch physicist, was able to show in a remarkable paper that the degree of coherence could be measured by studying the visibility of interference fringes. Later van Cittert and Zernike, independently, formulated a central theorem concerning the propagation of coherence for stationary sources [62, 63]. The essence of the theorem due to van Cittert and Zernike is that the two-point coherence function in the far field of a quasimonochromatic spatially incoherent light source is proportional to the Fourier transform of the source intensity distribution. This theorem has applications for long-baseline interferometry in astronomy, in which one uses a measurement of the coherence function to obtain the source intensity distribution. Zernike's great contribution, for which he won the Nobel prize in 1953, was the invention of the phase contrast microscope. The basic idea is to introduce a phase shift in the light scattered by a small object and to make it interfere with the light of a coherent background. In this way the phase difference is thus converted into an amplitude difference.

After the first third of the twentieth century several factors conspired to shift the center of gravity of polarization and coherence theories toward the United States. An important example came with Wiener. In 1930 Norbert Wiener, at the MIT Mathematics Department, developed a rigorous mathematical basis for the theory of the coherence of light. Although his analysis is rather complicated, he was able to derive a measure of coherence in his historic paper "Coherency matrix and quantum theory" [64]. Note that Wiener researches on generalized harmonic analysis were motivated partly by investigations in the theory of Brownian motion and partly to give a rigorous interpretation of white light. He was dissatisfied with the conventional description of unpolarized light because this ignores the statistical nature of light waves. What is seldom appreciated is that Wiener belongs to that small group of theoretical physicists who shaped modern coherence theory. Autocorrelation and cross-correlation functions between field variables at two space-time points were introduced to describe correlations of random processes in electromagnetic fields. Wiener, and a few years later Khintchine, formulated a fundamental theorem of the theory of random processes that is central in the analysis of optical coherence, specifically, that the Fourier transform of the

autocorrelation function of a stationary random process is the spectral density (power spectrum) of the process. Although Wiener gave real grounds for the concept of coherence, his ideas did not gain general recognition. To quote Levinson, “most of Wiener’s important work was inspired by physics or engineering and in this sense he was very much an applied mathematician. He formulated his theories in the framework of rigorous mathematics and as a consequence his impact on engineering was very much delayed.” The close relation to the mathematical description of polarization states with spinors was noted by Jordan [65] in 1927, thus indicating the possibility of using the wealth of mathematical techniques developed in quantum mechanics for the treatment of problems involving partially polarized electromagnetic radiation, although Ugo Fano (1949) is recognized as being the first to have popularized the connection between the quantum treatment of polarization and the classic description of polarization of light in his paper “Remarks on the classical and quantum-mechanical treatment of partial polarization” [66].

The credit for pioneering modern sheet polarizers belongs to Land (perhaps best known as an inventor for conceiving and developing commercial applications for instant photography as early as 1947) [67]. In 1927 Edwin H. Land, then a freshman (19-year-old) physics major at Harvard College, invented the first synthetic polarizers by orienting crystalline needles of herapathite in a sheet of plastic through electro-, magnetostatic, as well as mechanical forces (*J*-sheet polarizer). Note that the material—sulfate of iodoquinine—which was the starting point for Land’s invention of plastic sheet polarizer, was discovered by William Herapath in the mid-1800s [28]. The list of applications of these polarizers is extensive and we do not attempt to review it here [1]. In the years 1938–1939 Land invented *H* and *K* (with Howard Rogers) synthetic sheet polarizers that do not contain dichroic crystals but form molecular analogs of the wire-grid polarizer. These molecular polarizers made of polyvinyl alcohol stained with iodine for *H*-sheet and polyvinylene for *K*-sheet possess a good stability to light and heat, can be large in size, and can act over a wide angular range. The developments of these optical devices required several advances in material research; polymer science has played a major role in this story early on. Land took out more than 500 patents related to different areas of research including polarization, photography, and human color vision.

The development of polarization applications of light focused attention on the importance of investigating methods for describing the changes in polarization state that a beam of light undergoes on interaction with optical elements. Several remarkable contributions were carried out successively by Paul Soleillet, Robert Clark Jones, Francis Perrin, and Hans Mueller. Individually, each had important significance. Collectively, they heralded what we know as “Jones and Mueller polarization matrices.” Soleillet’s deep insight into this subject area is already evident in a remarkable (little-known) paper published in 1929, “Sur les paramètres caractérisant la polarisation partielle de la lumière dans les phénomènes de fluorescence” [68]. Less well known is the fact that he was the first to suggest how one might characterize an optical device by the linear

relation between the output Stokes parameters and the input Stokes parameters. Although the  $4 \times 4$  polarization matrix relating the output Stokes parameters to the input Stokes parameters is now traditionally termed the *Mueller formalism*, Paul Soleillet was evidently the first to consider it. His paper could have become the cornerstone on which optical scientists could base their theoretical studies. But apparently Soleillet's contribution did not attract the attention it deserves, and the matrix formalism was later named for Hans Mueller, who rediscovered the results that can be found in Soleillet's original paper. The power and fecundity of what is now termed the *Jones calculus* in polarization optics was demonstrated by Robert Clark Jones. In 1941/42, he had a remarkable year at Land's laboratory where, as an undergraduate, he proved his first results. The exceptional pertinence of Clark Jones can be appreciated in the reading of the set of eight publications noted for their clarity and insight, "A new calculus for the treatment of optical systems," on the problems related to the effect of various non-image-forming optical instruments on an incident polarized electromagnetic wave using the formalism of matrix calculus [69]. The state of polarization is represented by a  $2 \times 1$  vector and each optical element is represented by a  $2 \times 2$  matrix. His formalism was initially devised to study the effect of material anisotropy on the state of polarization of light and led to the identification of eight types of canonical behaviors associated with birefringence and dichroism. This matrix method, however, has limitations: the wave is normally incident, completely polarized, and the theory is valid under the paraxial approximation. Moreover, this method does not consider reflection from the surface of an anisotropic medium, or the effect of multiple reflections within an anisotropic plate. The Jones matrix method was later extended by Wolf to deal with partial polarization [70]. Further significant extensions and elaborations of Jones' work were subsequently published by Berreman and others and by Vernon and Huggins, to treat the problem of reflection by and transmission through a slab of inhomogeneous anisotropic material assuming polarized light obliquely incident on the slab [71, 72]. One notable contribution to the study of scattering of polarized light by optical media was the work of Francis Perrin [73]. In his 1942 landmark paper "Polarization of light scattered by isotropic opalescent media," he extended previous work of R. Krishnan [74] and noted that symmetry properties of the optical medium permit reduction of the number of independent parameters of polarization matrices. He recognized the importance of the concept of symmetry in elastic scattering by isotropic suspensions of particles and rigorously exploited the constraints required by symmetry. Later Perrin and Abragam reported several symmetry relations for the special cases of scattering by spherical particles and for molecular scattering [73]. The matrix method to represent the effect of non-image-forming optical instruments and scattering media on incident Stokes parameters is commonly termed the *Mueller formalism* after Hans Mueller, then a professor of physics at MIT. Apparently he had never seen Soleillet's original paper. Mueller used such matrices in lectures on optics at the MIT during 1945–1948 but did not publish

his results in the archival literature except in a short note: "The foundations of optics" [75]. An MIT report is cited by several authors, but I was not able to locate any other published works. Progress worth noting was also made by N. G. Parke, who investigated the connection between the Jones and Mueller matrix formalisms in his Ph.D work done at MIT under the guidance of N. Wiener [76].

On the experimental side, progress was rapid. As early as 1928, the French astronomer Bernard Lyot, gave the first impetus to the use of polarization in remote sensing of planetary surfaces [77]. In 1928 Lyot's paper on what is now called the *Lyot depolarizer* was published. This depolarizer consists of two retardation plates, with retardation in the ratio 1:2 and with their fast axes oriented at 45° to one another. He also invented several optical devices such as the double-refraction filter (1944). In the years 1934 and 1935 important papers by Langsdorf Jr., Wood, Birge, and DuBridge deserve special mention because they drew attention to the fact that the description of unpolarized light departs from the totally polarized case [78]. In 1947 Goos and Hänchen discovered experimentally the effect that now bears their names, related to the lateral spatial displacement of beams on reflection at interfaces between two dielectric media [79]. As earlier as 1943, Harold H. Hopkins demonstrated the possibility of controlling the spatial coherence of light for practical applications, such as microscopy [80]. After years of effort, ellipsometric measurements occurred that changed the study of surface and thin-film phenomena. Ellipsometry is a technique for *in situ* nondestructive characterization of surface (interfacial) phenomena using the change in the state of polarization of a lightwave probe. The term *ellipsometry* is fairly recent and was coined by Alexandre Rothen in 1945 [81]. The earliest applications of ellipsometry in biology were investigated by Rothen. It was von Frisch's achievement to show that many insects, such as bees, ants, and hornets are able of orientating themselves by using the polarization of sky light to find their way. The eyes of these invertebrates have birefringent properties and have the ability to sense the direction of polarization of skylight [82]. Because vision dominates human perception, it is quite surprising that the human visual system is almost completely insensitive to polarization. However, William Shurcliff called attention to the fact that the human eye has the ability to distinguish between circularly polarized from unpolarized light, in relation to the phenomenon of Haidinger's brush. Shurcliff is also noted for an important monograph on polarized light [94].

On the theory side, Luneberg developed, in 1944, an analytic formulation of diffraction theory taking into account polarization phenomena [83]. A list of classic highlights would not be complete without tribute to Henry Hurwitz. In 1945, an important paper by Hurwitz, "The statistical properties of unpolarized light," was published and attracted notice. Specifically, Hurwitz was able to compute the average values of many polarization parameters, including the ellipticity angle [84]. In 1946 the prominent physicist Denis Gabor invented the analytic signal concept for making a unique complex representation of a

real signal and used it to treat problems arising in communication theory. This representation is a generalization of one that is used for monochromatic radiation bringing to the field of optical coherence theory a thorough theoretical foundation [85]. Gabor's name is also on the lips of optical scientists for the invention of the technique of holography (the wavefront reconstruction process) in 1948 and for which he received a Nobel prize in physics for this breakthrough in 1971. The required physical basis in the theory of electromagnetic energy propagation in random media was suggested by Subrahmanyan Chandrasekhar, professor of astronomy and astrophysics at the University of Chicago in 1946, in his book *Radiative Transfer*. In this monograph, he used the Stokes parameters method in his solution of the problem of radiative transfer (Boltzmann equation approach) for the scattering of partially polarized light [86]. This work launched the theory of vector radiative transfer to model light propagation in anisotropically scattering media. Thanks to his profound scholarship, he clearly detected the trends in the development of modern physics and was awarded the Nobel prize in 1983 for his pioneering achievements in astrophysics. Another big advance in the study of radiation scattering by matter came when van de Hulst explained the polarization characteristics of the glory. He contributed much to the field by writing an important monograph, *Light Scattering by Small Particles*, emphasizing the symmetry constraints into the analysis of polarization matrices [87]. Other theoretical progress should be mentioned. To our knowledge, the first appearance in the optics literature, of a geometric algebra relying on quaternions to compute rotations and retardations of pure states of polarization was in a paper by Richartz and Hsü in 1949, for predicting the change of polarization by optical devices [88]. However, their formalism cannot describe the action of a polarizer and even partially polarized light.<sup>2</sup> In the years 1951 and 1952, the propagation of polarized light in optically active crystals was analyzed in a series of important articles of Ramachandran and Ramaseshan [89]. The application of the density matrix formalism in the description of phenomena connected with electron and photon polarization was discussed by Tolhoek [90]. The subject was taken up by several other contributors: Falkoff and MacDonald [91], Westfold [92], and independently Jerrard [93] developed a formalism based on the complex number representation (i.e., the ratio between the complex amplitudes of the components of the electric vector in an orthogonal coordinate system normal to the direction of wave propagation) to represent the state of polarization of polarized radiation transmitted through birefringent media.

Here we pause to make a brief comment. For many reasons the history of science in the former Soviet Union is not fully known in the West. For example, G. V. Rozenberg [95] and D. G. Stamov [96] were pioneers in developing the study of polarized light scattering in the atmosphere. These

<sup>2</sup>Recent work by Pellat-Finet has extended the mathematical formulation of the Richartz and Hsü quaternion method to deal with polarizers. See, for instance, P. Pellat-Finet, *Optik* **84**, 1990, 169; *ibid.* **90**, 1992, 101.

works were overlooked in countries other than Russia for several years, but their importance has since been appreciated.

During World War II, there was a substantial increase of interest and technological development of radar into the microwave range for aircraft detection. The bridges between pure science and military technologies that have been created in and for war continued in various forms during the postwar years. The wartime research laid the groundwork for the development of polarization radar. In the early 1950s, experimental efforts to make coherent polarimetric phase information the basis for technologically important devices a reality were undertaken. An influx of investigators with electrical engineering backgrounds led to work on polarization radar where the game is to guess the shape of an unknown scattering object from an analysis of the scattering data. One of the earlier pioneers in this area was Sinclair at Ohio State University [97]. The state of polarization of an electromagnetic wave scattered by a target differs, in general, from that of the incident wave. To characterize the scattering properties of the target, Sinclair introduced, in 1950, the  $2 \times 2$  coherent backscattering matrix formalism (the  $2 \times 2$  Sinclair matrix is analogous to the  $2 \times 2$  Jones matrix for the case of transmission). These efforts were pursued by Edward Morton Kennaugh, who introduced, in 1952, his radar target characteristic operator theory, based on the optimal polarization null concept [98]. He demonstrated that there exist polarization states for which the radar receives minimum or maximum power. Other works that have pushed this area of development very hard and need be cited are those of Deschamps and Huynen [99, 100]. More specifically, Jean Richard Huynen made a seminal contribution to polarimetric radar theory techniques by introducing the fork concept basis to study the polarimetric radar target optimization problem [100]. In subsequent years, various methods to study the optimal reception problem, namely, to find incident polarization states to optimize, for a given scattering or Mueller matrix system, the intensity at the receiver, were developed and large amounts of data were accumulated [101, 102]. These methods have also found application in such fields as geologic surveys and medical and remote sensing diagnostics.

During the years following the war, theoretical progress had impressive success in interpreting the statistical properties of light. A great step in coherence theory was taken by Emil Wolf during 1954/55 in his landmark papers "Optics in terms of observable quantities" and "Coherence properties of partially polarized electromagnetic radiation" giving a treatment of partial coherence, that has a great deal in common with the theory of partial polarization [70]. They must be considered along with the studies of Wiener as the classic papers in the field. The foundations of the modern polarization theory were laid. The starting point for most calculations having a bearing on optical coherence theory is the  $2 \times 2$  Wolf coherency matrix. His coherency matrix differs from that of Wiener in that Wolf chose to take as his fundamental quantities temporal covariance functions, unlike Wiener, who used integrated power spectra. Mathematically these two approaches are functionally

related through a Fourier–Stieljes transform. In a series of subsequent widely quoted papers and two important books coauthored by Max Born [103] (*Principles of Optics*) and Leonard Mandel [104] (*Optical Coherence and Quantum Optics*) noted for their clarity and insight, he provided a solid and transparent basis in terms of the physical interpretation for the theoretical understanding of the concept of coherence. Wolf also introduced more general three-dimensional tensors for dealing with nonplane waves [105]. Under his tutelage a large school of scientists in the field of optical physics emerged, and they continue to do important work to the present day. Among the numerous contributions of Leonard Mandel in the subject, we mention an analysis of coherence when the spectral width of the radiation is not narrow, and the introduction of the concept of cross-spectral purity [106]. Within this context, researchers in increasing numbers began to study optical coherence. These works have resulted in the explanation of a wealth of phenomena and the development of myriad new results; we refer the reader to Blanc-Lapierre and Dumontet [107], Gamo [108], and Parrent and Roman [110]. Bourret was the first to calculate the coherence properties of blackbody radiation in equilibrium in an enclosure that permit to characterize the spatial and temporal fluctuations of a (thermal) electromagnetic field [109]. Subsequently Richard Barakat added, in his 1963 paper “Theory of the coherency matrix for light of arbitrary spectral bandwidth,” the concept of spectral coherency matrix from a somewhat different point of view than the one introduced by Wiener [111]. Thus in the early 1960s, the conceptual basis of the coherency (density) matrix was thoroughly formulated.

Further important theoretical contributions to the subject were made by the great Indian physicist S. Pancharatnam and can be singled out. For example, in 1954, he introduced the concept of spectral functions to deal with the description of the polarization properties of a polychromatic light beam [112]. As early as 1956, in a classic paper, he was able to predict a remarkable effect, now usually called the *geometric phase*, which concerns the phase change of a light beam whose polarization state is made to trace out a cycle on the Poincaré sphere. He further observed this phase in a study devoted to the interference of two coherent light beams. Pancharatnam quickly realized that the origin of this phase is topological. It is remarkable that when Pancharatnam discovered this important effect, he was only 22 years of age. It is also worth to note that Pancharatnam belonged to the distinguished dynasty that includes the spectroscopist Raman, the astrophysicist Chandrasekhar, and the crystallographer Ramaseshan. In addition to being important in its own right, this discovery has stimulated many subsequent developments, including what is termed in the contemporary literature as Berry phase [112]. Here it is also worth mentioning the important contribution of Saxon on reciprocity and symmetry theorems that complements the earlier works of Rayleigh, Krishnan, and Perrin [113].

Advances in instrumentation came on many different fronts. Lyot and Öhman designed birefringent filters by means of which the distribution of

hydrogen in the sun may be measured by photographing the solar corona using H<sub>α</sub> line [114]. Solc and independently Evans discovered polarization interference filters composed of a stack of identical birefringent plates (each oriented at a prescribed azimuth angle) placed between crossed polarizers [115]. This discovery has found use an important application where filters of extremely narrow bandwidth with wide angular fields or tuning capability are required, including electrooptic tunable filters and electrically tunable lasers. Extensive polarization measurements in the area of astrophysics were pioneered by Hiltner and Dollfus and led to many important results concerning the remote sensing of planetary surfaces and atmospheric constituents [116]. In the 1950s considerable improvement in instrumentation focused attention on the importance of investigating higher-order-than-two correlation effects in optical fields. In the years 1952–1957, Hanbury-Brown and Twiss showed that when a lightwave illuminates two photoelectric detectors, the outputs from the detectors are in general correlated; this correlation is proportional to the square of the degree of coherence of the light vibrations at the detectors. They built the first optical interferometer to measure the apparent size of visible stars and performed successful experiments demonstrating that correlations between intensities (i.e., fourth-order effect in optical fields) can also be measured. These correlation phenomena were used to determine the velocity of light and the profiles of spectral lines [117]. The important point made by Hanbury-Brown and Twiss is that the degree of coherence at two points in a radiation field can be inferred from correlation measurements of the fluctuating signals appearing at radio antennas placed at the two points.

Certain periods of history bear the name of the materials that have shaped the development of civilization, such as the Stone age. An appropriate characterization of optics in the second half of the twentieth century would be the “laser age,” since the laser is hailed as one of the most significant inventions that has truly revolutionized the field. The *experimentum crucis* came in 1954 when Charles Townes and his coworkers realized the first maser, which is an acronym for microwave amplification by stimulated emission of radiation. Townes and coworkers designed a microwave oscillator based on stimulated emission by molecules (NH<sub>3</sub>) in an excited state [118]. Two years later, Nicolaas Bloembergen proposed the first tunable three-level solid-state maser [119]. Subsequently in 1960, Arthur Schawlow and Charles Townes extended the concept of maser to higher frequencies of the electromagnetic spectrum. *Infrared and Optical Masers*, based on a plane and parallel Fabry-Perot interferometer as optical resonator to provide a strong coupling between light and the amplifying laser materials [120]. This laid the groundwork for development of the laser. In recognition of their pioneering researches, Charles Townes shared the 1964 Nobel prize for physics with Nicolai Basov and Aleksandr Prokhorov for their fundamental work in quantum electronics that eventually led to the construction of oscillators and amplifiers based on the maser–laser principle. Note that Bloembergen and Schawlow will receive the Nobel prize for physics in 1981 for their contributions to the development of

laser spectroscopy. Then, Theodore H. Maiman realized the first laser using the fluorescence of a ruby crystal pumped by a xenon discharge lamp [121]. Shortly thereafter followed the operation of the first gas (helium-neon) laser by Ali Javan [122], and the following year Robert Hall and his coworkers described the first semiconductor laser [119]. By contrast with a usual thermal light source, such as an incandescent (tungsten) lamp or the heating element of a stove, a laser source produces coherent light amplified by stimulated emission in an active medium inside an optical resonator, whose spectrum is exceedingly narrow and that is highly directional. The historical development of lasers is well documented in Ref. 119. In physics, the laser has galvanized the science of the interaction of radiation with matter. Its imprint stamps a huge range of experimental techniques and has posed interesting new problems in the search to understand the nature of light; for instance, the fluctuation properties of such radiation cannot be described by a Gaussian random process. In technology it has resulted in expanded techniques and key contributions to communication systems, such as guiding dielectric structures.

In the more technical arena, a number of interesting devices were fabricated in the 1960s. George R. Bird and Maxfield Parrish Jr. designed and demonstrated wire-grid infrared polarizers [123]. A wire-grid polarizer consists of a grid of parallel metal wires that reflects one polarization (e.g., electric field parallel to the grid lines) of incident radiation field while transmitting the other (e.g., electric field perpendicular to the grid lines), provided the period of the grid is smaller than the wavelength of the wave. Measurements of the Mueller matrix were pioneered by Pritchard and Elliott [124] and Rozenberg [125] for characterization of the optical properties of the atmosphere. Inoué, Hyde, and Kubota carried out the theory of the diffraction image in the polarizing microscope, based on the fact that refraction at the optical surfaces of a lens system modifies the polarization of the wave transmitted through it [126].

In the early 1960s many developments culminated in the establishment of quantum optics, which gave the basic understanding of laser light. Quantization is necessary to show the bosonic nature of photons. In 1963 Roy J. Glauber of Harvard University introduced quantum analogs of the correlation functions of the classic theory as expectation values of normally ordered products of creation and annihilation operators of the electromagnetic field [127]. The quantum theory of coherence phenomena was originated mainly by Glauber. He gave the quantum state of the laser field its "official" name: the coherent state. Sudarshan [128] investigated the tradeoff between the classic and quantum-mechanical descriptions of coherence and found that quantum correlation functions take the same form as the classic versions if a particular positive definite phase-space function (termed the " $P$  distribution" by Glauber) is used for the description of the statistical properties of the field. Mandel and Wolf did much to develop the foundations of the quantum statistical properties of radiation [104]. Jauch and Rohrlich presented a quantum-mechanical approach to polarized light and introduced Stokes operators that contain all the relevant information about the polarization of the quantized

field [129]. Building on the base laid by quantum mechanics, photoelectron counting attracted new converts to the field and physicists formulated the theoretical foundations to study the interactions of the electromagnetic field with electrons [130].

The theoretical formulation of nonlinear optics was done by Nicolaas Bloembergen at Harvard University in 1956 and became a fundamental piece of work [131]. Following Bloembergen's work, researchers began to understand that nonlinear optical materials could have properties quite different from those of conventional linear materials. In 1961 the key discovery of generation of second harmonics was made by Franken and his colleagues. They detected UV (ultraviolet) light ( $0.347\text{ }\mu\text{m}$ ) at twice the frequency of a ruby laser beam ( $0.694\text{ }\mu\text{m}$ ) when this beam traversed a quartz crystal [132]. This key discovery was one of the earliest nonlinear optical phenomena to be observed experimentally. Maker, Terhune, and Savage were the first to demonstrate that a strong elliptically polarized pulse propagating within a nonlinear isotropic medium induces a refractive-index change that results in a continuous precession of the orientation angle of the polarization ellipse—the self-induced ellipse rotation—while leaving its shape and handedness unchanged [133]. The discovery of this optically induced birefringence was the spark that ignited the explosion of energy devoted to the nonlinear optical properties of materials. For example, optical parametric downconverters serve as tunable infrared coherent oscillators and microwave electrooptic modulators contribute to the development of optical fiber communication systems. But a discussion of these topics is beyond the scope of this book.

In this brief historical survey, we have spanned the exciting three-century period from Bartholinus's memoir of 1669 to the research report of Maker, Terhune, and Savage in 1964. A large number of physicists, astronomers, chemists, electrical engineers, and mathematicians have been involved in these efforts and their work has resulted in a vast edifice of scientific knowledge and technique that has led to many technological revolutions. Since an enormous volume of archival literature, both theoretical and experimental, which was concerned in the properties not only of polarized light but of other frequencies of the electromagnetic spectrum as well, has appeared. I have taken some trouble to sketch these highlights because the fundamental science involved in understanding polarization of light has great intellectual and practical value. The history has been enriched because of the many diverse fields, including optical communications, optical signal processing, spectroscopy, and imaging in which this body of knowledge has found engineering applications. A further enumeration of the current trends would be a formidable undertaking. Most of these developments have not been fully explored on either experimental or theoretical fronts and are much too recent to be written up as "history." But with all the important accomplishments, are there any things left for the future? Most definitely, yes. The last word on polarization optics has not yet been heard. There is very good reason to believe that with our increasingly powerful methods, we are going to find surprises not only in optical physics but in

technology as well. If we have gone so far in our understanding of the nature of polarized light in this century, it is in large part by standing on the shoulders of intellectual giants such as Poincaré, Faraday, and Maxwell. If Fresnel, Arago, and their colleagues had been told in 1800–1820 that they were laying the foundations for technological processes such as optical fiber communications and imaging of tumors for medical diagnostics, they would have been somewhat startled. I am confident that the future of polarization optics will be as exciting and fruitful as the past has been.

## Further Reading

Of several books that fell into my hands at an early stage in the manuscript, the following were particularly rewarding:

1. My favorite introductory text in this area is W. Shurcliff, *Polarized Light: Production and Use*, Harvard University Press, Cambridge, 1962.
2. J. Strong, *Concepts in Classical Optics*, Freeman, San Francisco, 1958.
3. M. Born and E. Wolf, *Principles of Optics*, 6th ed., Pergamon Press, Oxford, 1980.
4. See V. Ronchi, *Histoire de la Lumière*, Colin, Paris, 1957. French transl. J. Taton, for an account of the history of the development of the theory of light. See also A. Sabra, *Theories of Light from Descartes to Newton*, Oldbourne, London, 1967.
5. The work of the period 1820–1830 is skillfully summarized and critically discussed by J. Z. Buchwald, *The Rise of the Wave Theory of Light*, Univ. Chicago Press, Chicago, 1989. Additional information on the historical background may be found in E. Whittaker, *A History of the Theories of Aether and Electricity*, Harper & Row, New York, 1960, Vols. I, II. Further references will be found in: M. E. Skolnik, *Radar Handbook*, McGraw-Hill, New York, 1970.
6. G. Cantor, "The reception of the wave theory of light in Britain," in *Historical Studies in the Physical Sciences*, R. McCormach, ed., Princeton Univ. Press, Princeton, 1975, Vol. 6.
7. W. Swindell, ed., *Benchmark Papers in Optics, Polarized Light*, Dowden, Hutchinson, and Ross, Stroudsburg, 1975. A collection of reprints of 35 papers on polarized light selected from the literature on this subject of the last three centuries.
8. E. Mach, *The Principles of Physical Optics: A Historical and Philosophical Treatment*, Methuen, London, 1926.
9. W. T. Scott, *Am. J. Phys.* **31**, 1963, 819. This review provides a full bibliography on the evolution of the electromagnetic field concept.
10. R. Wehner, *Sci. Am.* **23**, 1976, 106. See also H. LaFay, *Natl. Geog.* **137**, 1970, 528.
11. Craig Bohren, personal communication. See also C. Roslund and C. Beckman, *Appl. Opt.* **33**, 1994, 4754.
12. E. Bartholinus, Hafniae, 1670, reprinted in Ref. 1-7.
13. C. Huyghens, *Traité de la Lumière*, Van der Aa, Leiden, 1690. We also refer to *Oeuvres Complètes de Christiaan Huyghens*, Société Hollandaise des Sciences, Nijhoff, La Hague, 1950 and to A. Ziggelaar, *Ann. Sci.* **37**, 1780, 179.

14. Sir Isaac Newton, *Opticks*, Dover Publications, 1952 (edition based on the 4th edition, Royal Society, London, 1704), pp. 107–111. See also A. E. Shapiro, ed., *The Optical Papers of Isaac Newton*, Cambridge Univ. Press, Cambridge, UK, 1994.
15. T. Young, *Phil. Trans. Roy. Soc. (London)* **92**, 1802, 387; *ibid.* **94**, 1804, 1. Interesting new aspects on Young's interference experiment are reviewed by D. James and E. Wolf, *Phys. Lett. A* **157**, 1991, 6.
16. E. Malus, *Mem. Soc. Arcueil* **1**, 1808, 113. See also C. F. Bohren, *Weatherwise*, 1988, 175 for interesting historical details about the discoveries of Malus.
17. D. B. Brewster, *Phil. Trans. Roy. Soc. (London)* **105**, 1815, 125. See also D. B. Brewster, *Phil. Trans. Roy. Soc. (London)* **105**, 1816, 60. For a biographical sketch of the life of Sir David Brewster, the interested reader is referred to E. S. Barr, *Opt. News* **16**, 1988, 8. A fascinating book about the invention of the kaleidoscope is D. B. Brewster, *The Kaleidoscope: Its History, Theory, and Construction*, reprinted by Van Cort Instruments, Holyoke, MA, 1987. See also A. Lakhtakia, *Opt. News* **15**, 1989, 14.
18. F. E. Neumann, *Berichte Konigl. Preuss. Akad. Wissensch.*, 1840.
19. A. J. Fresnel, *Oeuvres Complètes*, L'Imprimerie Impériale, Paris, 1866. See also A. J. Fresnel, *Ann. Chim. Phys.* **28**, 1825, 147.
20. F. Arago, *Oeuvres Complètes*, Gide, Paris, 1854–1858.
21. F. Arago and A. Fresnel, *Ann. Chim. Phys.* **2**, 1819, 288.
22. See, for instance, E. Collett, *Am. J. Phys.* **39**, 1971, 1483 and R. Barakat, *J. Opt. Soc. Am. A* **10**, 1993, 180.
23. J. W. Herschel, "Light," in *Encyclopedia Metropolitana* **4**, 1841, 341.
24. W. Nicol, *Edinburgh New Phil. J.* **6**, 1829, 83.
25. W. Haidinger, *Ann. Physik* **67**, 1846, 435. For an interesting discussion of the form and behavior of Haidinger's brushes, see also G. P. Misson, *Ophthalm. Phys. Opt.* **13**, 1993, 392 and L. L. Sloan, *J. Opt. Soc. Am.* **45**, 1955, 402. We also refer the reader to W. Haidinger, *Ann. Physik* **70**, 1847, 531.
26. W. A. Shurcliff, *J. Opt. Soc. Am.* **45**, 1955, 399. See also M. Minnaert, *Light and Color in the Open Air*, Dover, New York, 1954.
27. L. Pasteur, *Über die Asymmetrie bei Naturlich Vorkommenden Organischen Verbindungen*, Teubner, Leipzig, 1907. See also E. Charney, *The Molecular Basis of Optical Activity: Optical Rotatory Dispersion and Circular Dichroism*, Wiley, New York, 1979 for further historical details.
28. W. B. Herapath, *Phil. Mag.* **3**, 1852, 161; *ibid.* **9**, 1855, 366.
29. G. G. Stokes, *Trans. Camb. Phil. Soc.* **9**, 1852, 399. See also G. G. Stokes, *Camb. Dublin Math. J.* **4**, 1849, 1 and G. G. Stokes, *Mathematical and Physical Papers*, Cambridge Univ. Press, London, 1901.
30. G. Govi, *C. R. (Comptes Rendus) Acad. Sci. (Paris)* **51**, 1860, 360, 669.
31. J. Tyndall, *Proc. Roy. Soc. (London)* **17**, 1869, 223. See also J. Tyndall, *Philos. Mag.* **37**, 1869, 384.
32. M. Faraday, *Experimental Researches in Electricity: Faraday's Collected Works*, 3 vols., Dover, New York, 1964.
33. A. Zajonc, *Catching the Light: The Entwined History of Light and Mind*, Bantam, New York, 1993.

34. J. C. Maxwell, *Proc. Roy. Soc. (London)* **13**, 1864, 531. How familiar are many of today's scientists with the actual content of this paper beyond the customary ritual of using it as a reference? See also *The Scientific Letters and Papers of James Clerk Maxwell*, P. M. Harman, ed., Cambridge Univ. Press, New York, 1990, Vol. 1 and J. C. Maxwell, *A Treatise on Electricity and Magnetism*, 2 vols., Dover, New York, 1962. Another useful source is D. M. Siegel, *Innovation in Maxwell's Electromagnetic Theory*, Cambridge Univ. Press, Cambridge, UK, 1991.
35. A fuller account of the works of Maxwell and Faraday may be found in L. Rosenfeld, *Nuovo Cimento Suppl.* **4**, 1957, 1630 and P. M. Harman, *Eur. J. Phys.* **14**, 1993, 148.
36. T. S. Kuhn, *The Structure of Scientific Revolution*, Univ. Chicago Press, Chicago, 1970. See also P. Hoymingen-Huene, *Reconstructing Scientific Revolutions: Thomas S. Kuhn's Philosophy of Science*, Univ. Chicago Press, Chicago, 1993.
37. R. Kohlrausch and W. Weber, *Pogg. Ann. Physik. Chem.* **99**, 1856, 10.
38. E. Verdet, *Leçons d'Optique Physique*, Imprimerie Impériale, Paris, 1869. It appears to me that the paper E. Verdet, *Ann. Sci. Ecole Normale Supérieure* **2**, 1865, 291 represents a major involvement with the subject of unpolarized light.
39. J. Stefan, *Wien. Ber.* **50**, 380, 1864. See also J. Stefan, *Pogg. Ann.* **124**, 623, 1864.
40. J. Kerr, *Phil. Mag.* **50**, 1875, 337. See also A. Yariv and P. Yeh, *Optical Waves in Crystals*, Wiley, New York, 1984.
41. O. Wiener, *Ann. Phys.* **40**, 1890, 203.
42. H. Hertz, *Wiedem. Ann.* **34**, 1888, 551 and H. Hertz, *Miscellaneous Papers*, Macmillan, London, 1896.
43. J. F. Mulligan, *Am. J. Phys.* **55**, 1987, 711 and J. Hertz, M. Hertz, and C. Susskind, eds., *Heinrich Hertz: Memoirs, Letters, Diaries*, San Francisco Press, San Francisco, 1977 for biographical details.
44. L. V. Lorenz, *K. Dan. Vidensk. Selsk. Forh.* **6**, 1890, 1. See also H. Kragh, *Appl. Opt.* **30**, 1991, 4688 for interesting details concerning the scientific achievements of Lorenz. Note that a survey of the work of Lorenz and other early workers concerned with the study of the scattering of plane waves by a sphere was done by N. A. Logan, *Proc. IEEE*, **1965**, 773.
45. H. Poincaré, *Théorie Mathématique de la Lumière*, Georges Carré, Paris, 1892, Vol. 2.
46. J. W. Strutt (Lord Rayleigh), *Collected Scientific Papers*, Cambridge Univ. Press, Cambridge, UK, 1920; reprinted by Dover, 1964. See also J. W. Strutt (Lord Rayleigh), *Phil. Mag.* **41**, 1871, 107, 274, 447; *ibid.* **12**, 1881, 81; *ibid.* **49**, 1900, 157. For an entry into the enormous body of works made by Lord Rayleigh, see J. N. Howard et al., *Appl. Opt.* **3**, 1964, 1091 and A. Young, *Phys. Today* **35**, 1982, 2.
47. G. Mie, *Ann. Phys. (Leipzig)* **25**, 1908, 376. Apparently the importance of this 69-page paper appears to have been greatly underestimated by both its author and contemporary scientists. See also N. A. Logan, *J. Opt. Soc. Am.* **52**, 1962, 342 for interesting details on the Mie solution and P. Lilienfeld, *Appl. Opt.* **30**, 1991, 4696 for a biographic note.
48. P. Debye, *Ann. Phys.* **30**, 1909, 59.
49. Q. Majorana, *C. R. Acad. Sci. (Paris)* **135**, 1902, 159.
50. C. Mauguin, *Bull. Soc. Fr. Mineral.* **34**, 1911, 71.

51. A. Schuster, *Astrophys. J.* **21**, 1905, 1.
52. Historical developments in understanding the polarization of light in the sky are reviewed in G. V. Rozenberg, *Sov. Phys. Usp.* **3**, 1960, 346.
53. A. Cotton, *C. R. Acad. Sci. (Paris)* **120**, 1895, 989 and G. Bruhat, *Rev. Opt.* **8**, 1929, 365, 413.
54. A. A. Michelson, *Phil. Mag.* **21**, 1911, 554. See also G. P. Können, *Polarized Light in Nature*, Cambridge Univ. Press, Cambridge, UK, 1985 and G. W. Kattawar, *Opt. Phot. News* **6**, 1994, 43.
55. W. Roentgen, *Sitzungber. Phys.-Med. Gesellschaft Würzburg* **132**, 1895 (Engl. transl. in *Nature* **53**, 1896, 274). For biographic details, see also O. Glasser, *Wilhelm Conrad Roentgen and the Early History of X-rays*, Thomas, Baltimore, 1934 and W. R. Nitske, *The Life of Wilhelm Conrad Roentgen, Discoverer of the X Ray*, Univ. Arizona Press, Tucson, 1971.
56. C. G. Barkla, *Proc. Roy. Soc. (London) Ser. A* **77**, 1906, 247.
57. M. Planck, *The Theory of Heat Radiation*, Dover, New York, 1959. See also A. Ore, *Phys. Rev.* **98**, 1955, 887 and P. Rosen, *Phys. Rev.* **96**, 1954, 555 for details about the entropy of a polarized “pencil of radiation.”
58. M. von Laue, *Ann. Phys.* **20**, 1906, 365; *ibid.* **23**, 1907, 1, 795.
59. M. Berek, *Z. Physik* **36**, 1926, 675; *ibid.* **36**, 1926, 824; *ibid.* **37**, 1926, 387; *ibid.* **40**, 1927, 420. See also R. Wartmann, E. Wolff-Fischer, and M. Berek, *Leitz Mitt. Wiss. Techn.* **9**, 1986, 7.
60. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton Univ. Press, Princeton, 1955. See also J. von Neumann, *Collected Works*, A. H. Traub, ed., Pergamon Press, New York, 1961.
61. C. Lakeman and J. T. Grosmuller, *Physica* **8**, 1928, 193, 199, 305.
62. P. H. van Cittert, *Physica* **1**, 1934, 201; *ibid.* **6**, 1939, 1129.
63. F. Zernike, *Physica* **5**, 1938, 785.
64. N. Wiener, *Acta Math.* **55**, 1930, 117. See also N. Wiener, *J. Franklin Inst.* **207**, 1929, 525 and N. Wiener, *J. Opt. Soc. Am.* **43**, 1953, 225 and N. Wiener, *The Fourier Integral and Certain of Its Applications*, Cambridge Univ. Press, New York, 1933. We also refer to the special issue of the *Bulletin of the American Mathematical Society* (**72**, 1966, 1) dedicated to the memory of Norbert Wiener in recognition of his pioneering contributions to science. See also A. Khintchine, *Math. Ann.* **109**, 1934, 604.
65. P. Jordan, *Z. Physik* **44**, 1927, 292.
66. U. Fano, *J. Opt. Soc. Am.* **39**, 1949, 859. See also U. Fano, *Phys. Rev.* **93**, 1954, 121 and U. Fano, *Rev. Mod. Phys.* **29**, 1957, 74.
67. E. H. Land, *J. Opt. Soc. Am.* **41**, 1951, 957. See also E. H. Land and C. D. West, *Colloid Chem.* **6**, 1946, 160 for details about the discovery of dichroic polarizers. A reader wishing to appreciate the numerous patents by Land is advised to consult the annotated bibliography prepared by Schurcliff (Ref. 94). The interested reader is also referred to the special issue of *Optics and Photonics News* (**5**, 1994, 9) dedicated to the memory of Edwin Land in recognition of his pioneering contributions to science and technology.
68. P. Soleillet, *Ann. Physique* **12**, 1929, 23.
69. R. Clark Jones, *J. Opt. Soc. Am.* **31**, 1941, 488, 493, 500; *ibid.* **32**, 1942, 486; *ibid.* **37**, 1947, 107, 110; *ibid.* **38**, 1948, 671; *ibid.* **46**, 1956, 126.

70. E. Wolf, *Proc. Roy. Soc. A* **225**, 1954, 96. See also E. Wolf, *Nuovo Cimento* **12**, 1954, 884; E. Wolf, *Proc. Roy. Soc. A* **230**, 1955, 96; E. Wolf, *Phil. Mag.* **2**, 1957, 351; and E. Wolf, *Nuovo Cimento* **13**, 1959, 1165.
71. R. J. Vernon and B. D. Huggins, *J. Opt. Soc. Am.* **70**, 1970, 1364.
72. D. W. Berreman, *J. Opt. Soc. Am.* **62**, 1972, 502. See also P. J. Ling and S. Teitler, *J. Opt. Soc. Am. A* **1**, 1984, 703 and S. Teitler and B. W. Henvis, *J. Opt. Soc. Am.* **60**, 1970, 830.
73. F. Perrin, *J. Chem. Phys.* **10**, 1942, 415; F. Perrin and A. Abragam, *J. Phys. Rad.* **12**, 1951, 69.
74. R. S. Krishnan, *Proc. Indian Acad. Sci. A* **9**, 1938, 21; *ibid.*, 1938, 91; *ibid.*, 1938, 98; R. S. Krishnan, *Kolloid Zeits.* **84**, 1939, 2; *ibid.* **84**, 1938, 18 and R. S. Krishnan, *Proc. Indian Acad. Sci. A* **10**, 1939, 395.
75. H. Mueller, *J. Opt. Soc. Am.* **38**, 1948, 661. R. C. Jones also cited a declassified report by H. Mueller, report No. 2 of OSRD, Project OEMsr-576, 1943. However, the author has been unable to locate this report in the collections of the libraries to which he had access.
76. N. G. Parke III, *J. Math. Phys. MIT* **28**, 1949, 131.
77. B. Lyot, *Ann. Observ. Paris* **8**, 1928, 102. See also B. Lyot, *Ann. Astrophys.* **7**, 1949, 31. For additional details, see W. G. Egan, *Photometry and Polarization in Remote Sensing*, Elsevier, New York, 1985, B. H. Billings, *J. Opt. Soc. Am.* **41**, 1951, 966; A. P. Loebel, *J. Opt. Soc. Am.* **72**, 1982, 650, and P. H. Richter, *J. Opt. Soc. Am.* **69**, 1979, 460.
78. A. Langsdorf, Jr. and L. A. Du Bridge, *J. Opt. Soc. Am.* **24**, 1934, 1. See also R. W. Wood, *J. Opt. Soc. Am.* **24**, 1934, 4; R. T. Birge, *J. Opt. Soc. Am.* **25**, 1935, 179 and L. A. Du Bridge, *J. Opt. Soc. Am.* **25**, 1935, 182. We also refer to P. Hariharan and R. G. Singh, *J. Opt. Soc. Am.* **51**, 1961, 1148 for the detailed description of improvements in the experiments of the above authors.
79. F. Goos and H. Hänchen, *Ann. Phys.* **6**, 1947, 333; *ibid. Ann. Phys.* **6**, 1949, 251.
80. H. H. Hopkins, *Proc. Phys. Soc. London* **55**, 1943, 116. See also H. H. Hopkins, *Proc. Roy. Soc. A* **208**, 1951, 263, *ibid.* **217**, 1953, 408 and H. H. Hopkins, *J. Opt. Soc. Am.* **47**, 1957, 508.
81. A. Rothen, *Rev. Sci. Instrum.* **16**, 1945, 26. Historical reviews on ellipsometry have been published by A. Rothen, in *Ellipsometry in the Measurement of Surfaces and Thin Films*, E. Passaglia, R. R. Stromberg, and J. Kruger, eds., Natl. Bur. Std. Misc. Publ. 256, U.S. Govt. Printing Office, Washington, DC, 1964 and A. C. Hall, *Surf. Sci.* **16**, 1969.
82. K. von Frisch, *Bees: Their Vision, Chemical Senses and Language*, Cornell Univ. Press, Ithaca, NY, 1950. A more recent discussion of polarized-light navigation by insects, fish, and mammals has been given by R. Wehner, *Sci. Am.* **235**, 1976, 106. See also G. D. Bernard and R. Wehner, *Vision Res.* **17**, 1019, 1977; S. Rossel and R. Wehner, *Nature* **323**, 128, 1969; C. W. Hawryshyn, *Am. Sci.* **80**, 164, 1992; and D. A. Cameron and E. N. Pugh, *Nature* **353**, 161, 1991.
83. R. K. Luneberg, *Mathematical Theory of Optics*, Brown Univ., Providence, RI, 1944.
84. H. Hurwitz, *J. Opt. Soc. Am.* **35**, 1945, 525.
85. D. Gabor, *J. Inst. Elect. Eng. (London)* **93**, 1946, 429.

86. S. Chandrasekhar, *Astrophys. J.* **104**, 1946, 110. See also S. Chandrasekhar, *Radiative Transfer*, Clarendon Press, Oxford, 1950 and S. Chandrasekhar and D. Elbert, *Trans. Am. Phil. Soc.* **44**, 1954, 643.
87. H. C. van de Hulst, *Light Scattering by Small Particles*, Wiley, New York, 1957.
88. M. Richartz and H. Y. Hsü, *J. Opt. Soc. Am.* **39**, 1949, 136.
89. G. N. Ramachandran and V. Chandrasekharan, *Proc. Indian Acad. Sci. A* **33**, 1951, 199. See also R. Ramaseshan, *Proc. Indian Acad. Sci. A* **34**, 1951, 32; G. N. Ramachandran and S. Ramaseshan, "Crystal optics," in *Handbuch der Physik*, S. Flügge ed., Springer-Verlag, Berlin, 1961, Vol. XXV/1, and G. N. Ramachandran and S. Ramaseshan, *J. Opt. Soc. Am.* **42**, 1952, 49.
90. H. A. Tolhoek, *Rev. Mod. Phys.* **28**, 1956, 277.
91. D. L. Falkoff and J. E. McDonald, *J. Opt. Soc. Am.* **410**, 1951, 861.
92. K. C. Westfold, *J. Opt. Soc. Am.* **49**, 1959, 717.
93. H. G. Jerrard, *J. Opt. Soc. Am.* **44**, 1954, 634.
94. W. A. Shurcliff, *Polarized Light and Optical Measurement*, Pergamon Press, Braunschweig, 1971.
95. G. V. Rozenberg, doctoral dissertation, Moscow Univ., 1946. See also G. V. Rozenberg, *Usp. Fiz. Nauk* **71**, 1960, 173.
96. D. G. Stamov, doctoral dissertation, Moscow Univ., Moscow, 1953.
97. G. Sinclair, *Proc. IRE* **38**, 1950, 148.
98. E. M. Kennaugh, "Effects of type of polarization on echo characteristics," Antennas Laboratory, The Ohio State University, Columbus, 1949; see also E. M. Kennaugh, "Polarization properties of radar reflections," M.Sc. thesis, The Ohio State University, Columbus, 1952.
99. G. A. Deschamps, *Proc. IRE* **39**, 1951, 540. See also G. A. Deschamps and P. E. Mast, *IEEE Trans. AP-21*, 1973, 474.
100. J. R. Huynen, "Phenomenological theory of radar targets," Ph.D. dissertation, Delft Univ. Technology, Delft, The Netherlands, 1970. See also dissertation update in P. L. E. Uslenghi, ed., *Electromagnetic Scattering*, Academic Press, New York, 1978.
101. C. D. Graves, *Proc. IRE* **44**, 1956, 248.
102. For an historical account in the field of radar polarimetry, we refer the reader to W. M. Boerner and Wei-Ling Yan, in *Electromagnetic Modelling and Measurements for Analysis and Synthetic Problems*, B de Neumann, ed., Kluwer, Amsterdam, 1991.
103. M. Born and E. Wolf, *Principles of Optics*, 5th ed., Pergamon, Oxford, 1975 is among the most frequently cited references in the optics research literature.
104. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*, Cambridge Univ. Press, New York, 1995. See also L. Mandel and E. Wolf, *Rev. Mod. Phys.* 1965.
105. P. Roman and E. Wolf, *Nuovo Cimento* **17**, 1960, 462, 477.
106. L. Mandel, *J. Opt. Soc. Am.* **51**, 1961, 1342.
107. A. Blanc-Lapierre and P. Dumontet, *C. R. Acad. Sci. (Paris)* **238**, 1954, 1005. See also A. Blanc-Lapierre and P. Dumontet, *Rev. Opt.* **34**, 1955, 1.
108. H. Gamo, *J. Appl. Phys. (Japan)* **25**, 1956, 431; *ibid.* **26**, 1957, 414. See also H. Gamo, *J. Opt. Soc. Am.* **47**, 1957, 976.

109. R. C. Bourret, *Nuovo Cimento* **18**, 1960, 347.
110. G. B. Parrent, Jr. and P. Roman, *Nuovo Cimento* **15**, 1960, 370. See also G. B. Parrent, Jr., *J. Opt. Soc. Am.* **49**, 1959, 787 and G. B. Parrent, Jr., *Optica Acta* **6**, 1959, 285.
111. R. Barakat, *J. Opt. Soc. Am.* **53**, 1963, 317.
112. S. Pancharatnam, *Proc. Indian Acad. Sci.* **A44**, 1956, 247, 398. See also S. Pancharatnam, *Proc. Indian Acad. Sci.* **57**, 1963, 218, 231 and S. Pancharatnam, *Collected Works*, Oxford Univ. Press, Oxford, 1975. For illuminating comments on Pancharatnam's paper, see also M. V. Berry, *J. Mod. Opt.* **34**, 1987, 1401.
113. D. S. Saxon, *Phys. Rev.* **100**, 1955, 1771.
114. B. Lyot, *C. R. Acad. Sci. (Paris)*, **197**, 1933, 1593 and B. Lyot, *Ann. Astrophys.* **7**, 1944, 31. See also Y. Öhman, *Nature* **41**, 1938, 157, 291 and Y. Öhmann, *Ark. Astron.* **2**, 1958, 165.
115. I. Solc, *Cesk. Cas. Fys.* **3**, 1953, 366; *ibid.* **10**, 1960, 16. See also I. Solc, *J. Opt. Soc. Am.* **55**, 1965, 621. We also refer the reader to J. W. Evans, *J. Opt. Soc. Am.* **48**, 1958, 142 and J. W. Evans, *Appl. Opt.* **2**, 1963, 193.
116. W. A. Hiltner, *Astrophys. J.* **106**, 1947, 231; *ibid.* **109**, 1949, 471; *ibid.* **114**, 1951, 241 and A. Dollfus, *Ann. Astrophys.* **19**, 1956, 83. We also refer the reader to J. Lenoble, *J. Phys. Radium* **18**, 1957, 47.
117. R. Hanbury-Brown and R. Q. Twiss, *Nature* **177**, 1956, 27; **178**, 1956, 1046, 1447. See also R. Hanbury-Brown, *Boffin*, IOP, London, 1991 and R. Hanbury-Brown, *The Intensity Interferometer*, Taylor & Francis, London, 1974. For important applications, we refer the reader to J. H. Sanders, *Nature* **183**, 1959, 312 and A. T. Forrester, *J. Opt. Soc. Am.* **51**, 1961, 253.
118. J. P. Gordon, H. J. Zeiger, and C. H. Townes, *Phys. Rev.* **95**, 1954, 282; *ibid.* **99**, 1955, 1164.
119. For an entry in the enormous body of works in the area of lasers and masers, we refer the reader to M. Bertolotti, *Masers and Lasers: A Historical Approach*, Adam Hilger, Bristol, 1983. For an historical overview of the development of lasers, see also J. Hecht, *Laser Pioneers*, Academic Press, New York, 1992.
120. A. L. Schawlow and C. H. Townes, *Phys. Rev.* **112**, 1958, 1940.
121. T. H. Maiman, *Nature* **187**, 1960, 493. See also T. H. Maiman, *Br. Commun. Electron.* **7**, 1960, 674.
122. A. Javan, W. R. Bennett, and D. R. Herriot, *Phys. Rev. Lett.* **6**, 1961, 106. See also A. Javan, "Gaseous optical masers," in *Les Houches Lectures 1964*, C. de Witt, A. Blandin, C. Cohen-Tannoudji, eds., Gordon & Breach, New York, 1965.
123. G. R. Bird and M. Parrish, Jr., *J. Opt. Soc. Am.* **50**, 1960, 886. See also G. R. Bird and W. A. Shurcliff, *J. Opt. Soc. Am.* **49**, 1959, 235.
124. B. S. Pritchard and W. G. Elliott, *J. Opt. Soc. Am.* **50**, 1960, 191.
125. G. V. Rozenberg, *Adv. Phys. Sci. USSR* **11**, 1968, 353. See also G. I. Gorchakov and G. V. Rozenberg, *Izv. Acad. Sci. USSR Atmos. Oceanic Phys.* **3**, 1967. For many reasons, the history of science in the former Soviet Union is not fully known in the West. For instance, the important book (in Russian) by A. V. Rzhanov et al., *Principles of Ellipsometry*, Nauka, Novosibirsk, 1979 is virtually unknown in the West.

126. S. Inoué and W. L. Hyde, *J. Opt. Soc. Am.* **46**, 1956, 372. See also H. Kubota and S. Inoué, *J. Opt. Soc. Am.* **49**, 1959, 191 and H. Kubota and H. Saito, *J. Opt. Soc. Am.* **50**, 1960, 1020 and W. C. McCrone, L. B. McCrone, J. G. Delly, *Polarized Light Microscopy*, Ann Arbor Science Publ., Ann Arbor, MI, 1978.
127. R. J. Glauber, *Phys. Rev.* **130**, 1963, 2529. See also R. J. Glauber, *ibid.* **131**, 1963, 2766 and R. J. Glauber, "Optical coherence and photon statistics," in *Les Houches Lectures 1964*, C. de Witt, A. Blandin, and C. Cohen-Tannoudji, eds., Gordon & Breach, New York, 1965.
128. E. C. G. Sudarshan, *Phys. Rev. Lett.* **10**, 1963, 277.
129. J. M. Jauch and F. Rohrlich, *Theory of Photons and Electrons*, Addison-Wesley, Reading, MA, 1955.
130. For an overview of the subject of photoelectron counting, see, for example, R. Barakat and J. Blake, *Phys. Rep.* **60**, 1980, 225 and references cited therein.
131. N. Bloembergen, "Nonlinear optics," in *Les Houches Lectures 1964*, C. de Witt, A. Blandin, and C. Cohen-Tannoudji, eds., Gordon & Breach, New York, 1965.
132. P. Franken, A. E. Hill, C. W. Peters, and G. Weinreich, *Phys. Rev. Lett.* **7**, 1961, 118.
133. P. D. Maker, R. W. Terhune, and C. M. Savage, *Phys. Rev. Lett.* **12**, 1964, 507. See also G. Mayer and F. Gires, *C. R. Acad. Sci.* **258**, 1964, 2039 and G. A. Askar'yan, *Sov. Phys. JETP* **15**, 1962, 1088.

---

**PART 2**

---

## **PRELIMINARIES TO A CLASSIC RADIATION FIELD THEORY**

Having chronicled in Part 1 the high points in the historical survey of polarization optics since its earliest days, we now turn our attention to a short introduction to Maxwell's theory of electrodynamics and the central ideas that underly all electromagnetic waves. To understand the many facets of light polarization, one must understand the fundamentals of the classic theory of electromagnetic radiation and its interaction with matter. This is the purpose of the present part of the book. The goal of our discussion is expository and tutorial. If this introduction sounds unfamiliar, the reader might wish to review the material in Refs. 1–7 for further information on this subject.

The presentation is as follows. Section 2.1 contains a description of the basic differential equations in the time-dependent form. I summarize several relevant definitions that should be found useful for understanding material in the remainder of this book and derive Fresnel's equations, which characterize the reflection and transmission of a plane electromagnetic wave at the plane boundary between two homogeneous isotropic media. This section also endeavors to explain physically, by reference to Wiener's experiment, why the electric vector is chosen to define the state of polarization of light. Then in Section 2.2 we encounter the constraints due to the invariance transformations. Section 2.3 is devoted to a discussion of the particular properties of monochromatic plane waves.

## SECTION 2.1

# The Basic Differential Equations and Boundary Conditions

### 2.1.1. MAXWELL'S EQUATIONS

Maxwell's equations describe the temporal and spatial dependence of electromagnetic fields. They give very good agreement with all experimentally observed classic phenomena to date, over a very large range in frequency, from dc to X rays. For lengths and times on an atomic scale the Maxwell equations must be combined with quantum theory.<sup>1</sup> However, for describing the effects of electromagnetic fields over distance large compared with atomic dimensions, the classic form of Maxwell's equations is reliable. The theory of electromagnetic processes rests on a set of linear equations (i.e., Maxwell's equations together with a set of material equations, which may not be linear), that express the response of matter to electromagnetic excitation. However, in the description of classic optical processes, they are linearized. Thus Maxwell's theory of electrodynamics is a linear theory. In the discussion below I use the terms "wave" and "light" interchangeably.

The following first-order, coupled partial differential equations relating the electric and magnetic fields in a linear, homogeneous and isotropic, medium (which will be assumed to fill the whole space) with sources are Maxwell field equations:

$$\operatorname{div} \mathbf{D}(\mathbf{R}, t) = \nabla \cdot \mathbf{D}(\mathbf{R}, t) = \rho \quad (2.1.1a)$$

$$\operatorname{div} \mathbf{B}(\mathbf{R}, t) = \nabla \cdot \mathbf{B}(\mathbf{R}, t) = 0 \quad (2.1.1b)$$

$$\operatorname{rot} \mathbf{E}(\mathbf{R}, t) = \nabla \times \mathbf{E}(\mathbf{R}, t) = -\frac{\partial \mathbf{B}(\mathbf{R}, t)}{\partial t} \quad (2.1.1c)$$

$$\operatorname{rot} \mathbf{H}(\mathbf{R}, t) = \nabla \times \mathbf{H}(\mathbf{R}, t) = \frac{\partial \mathbf{D}(\mathbf{R}, t)}{\partial t} + \mathbf{j}(\mathbf{R}, t) \quad (2.1.1d)$$

where the dot and the  $\times$  refer to scalar and cross-products, respectively. The notational scheme that I adhere to is as follows: a bold letter refers to a vector

<sup>1</sup>Maxwell's equations provide us a macroscopic theory. Now a coherent microscopic theory requires the quantum description of atoms and molecules. Quantum electrodynamics couples the free-space electromagnetic fields to the quantum-mechanical fields that characterize matter.

or a tensor quantity; a standard letter refers to a scalar quantity. I adopt the SI system of units.<sup>2</sup> The reader is warned that other systems of units, such as Gaussian cgs (centimeter–gram–second) units, lead to differences in the appearance of equations that can be confusing.

Here  $\mathbf{E}(\mathbf{R}, t)$  is the electric field ( $\text{V m}^{-1}$ ),  $\mathbf{H}(\mathbf{R}, t)$  is the magnetic field ( $\text{A m}^{-1}$ ),  $\mathbf{D}(\mathbf{R}, t)$  is the electric displacement ( $\text{C m}^{-2}$ ), and  $\mathbf{B}(\mathbf{R}, t)$  is the magnetic induction ( $\text{T}$ ), at a position vector  $\mathbf{R}$  in space and at time  $t$ . The term  $\mathbf{j}(\mathbf{R}, t)$  is the current density ( $\text{A m}^{-2}$ ) and  $\rho(\mathbf{R}, t)$  is the volume charge density ( $\text{C m}^{-3}$ ). In free space,  $\mathbf{E}$  and  $\mathbf{D}$  vary only by a multiplicative universal constant as to  $\mathbf{B}$  and  $\mathbf{H}$ . But in matter,  $\mathbf{D}$  includes the consequences of the electric polarizations of matter where  $\mathbf{H}$  includes the consequences of the magnetic polarization.

Maxwell's equations specify the behavior of the electromagnetic field in a neighborhood of a point  $\mathbf{R}$  and at time  $t$ . The set of these four differential equations completely describes the propagation of a radiation field in any medium. Linearity of Eqs. (2.1.1a–d) implies that the fields arising from several sources can be obtained by superposing the fields arising from each source treated independently.

To describe the interaction of light with matter at thermal equilibrium, Maxwell's equations need to be supplemented with the material equations<sup>3</sup>

$$\mathbf{D} = \varepsilon \varepsilon_0 \mathbf{E} \quad (2.1.2a)$$

$$\mathbf{B} = \mu \mu_0 \mathbf{H} \quad (2.1.2b)$$

$$\mathbf{j} = \sigma \mathbf{E} \quad (2.1.3)$$

where  $\varepsilon$  = the relative permittivity of the medium

$\mu$  = relative permeability of the medium

$\mu_0$  = permeability of vacuum =  $4\pi 10^{-7} \text{ H m}^{-1}$

$\varepsilon_0$  = permittivity of vacuum =  $1/\mu_0 c^2 = 8.85 10^{-12} \text{ F m}^{-1}$

$c$  = speed of light in vacuum =  $3 10^8 \text{ m s}^{-1}$

$\sigma$  = conductivity of the medium ( $\Omega^{-1} \text{ m}^{-1}$ ).

<sup>2</sup>A note about units—the Système International d'Unités, with the international abbreviation SI, is founded on seven base units: meter (m), kilogram (kg), second (s), ampere (A), kelvin (K), mole (mol), and candela (cd). All other units are derived units. Examples of SI derived units include coulomb (C), volt (V), henry (H), and farad (F).

<sup>3</sup>Note that Eqs. (2.1.2a,b) can be generalized to include optical rotation tensors. For plane-wave propagation in a homogeneous medium, the following material equation for an optically active material results:  $\mathbf{D} = \varepsilon \varepsilon_0 \mathbf{E} + i\varepsilon_0 \mathbf{G} \times \mathbf{E}$ , where  $\mathbf{G}$  is the gyration vector. For example, in the Faraday effect, which causes a rotation of the direction of vibration of light with distance when the material is placed in a magnetic field, we have  $\mathbf{G} = \gamma \mathbf{B}$ , where  $\gamma$  denotes the magnetogyration coefficient of the medium. The reader is referred to Refs. 2 and 4 for more details. Further elaborations of the theory concerning the most general linear complex media, the bianisotropic media, have considered the following material equations:  $\mathbf{D} = \varepsilon \varepsilon_0 \mathbf{E} + \zeta \mathbf{H}$  and  $\mathbf{B} = \mu \mu_0 \mathbf{E} + \xi \mathbf{H}$ , where the permittivity tensor  $\varepsilon$ , the permeability tensor  $\mu$ , and the cross-coupling tensors  $\zeta$ ,  $\xi$  are  $3 \times 3$  matrices.

Relations (2.1.2a–b) are consistent with the fact that the magnetic induction  $\mathbf{B}$  plays a role in magnetic phenomena analogous to that of the displacement vector  $\mathbf{D}$  in electrical phenomena.

I will restrict the discussion to the case for which  $\epsilon$ ,  $\mu$ , and  $\sigma$  are independent of the field strengths  $\mathbf{E}$  and  $\mathbf{H}$ . Quantities  $\epsilon$ ,  $\mu$ , and  $\sigma$  depend on many factors, notably on the temperature and frequency (dispersive medium). For an isotropic and homogeneous material medium,  $\epsilon$  and  $\mu$  are scalars independent of position that in general are complex quantities. If the material is not isotropic (e.g., crystalline media),  $\epsilon$  and  $\mu$  are non-Hermitian tensors of rank 2, within the general case of nine independent nonzero complex elements. The anisotropy may be natural or induced by some external source such as an electromagnetic field, or stress. As a general comment, we note that the frequencies in which we are interested ( $\sim 10^{15}$  Hz) are usually so high that the magnetic response of spins inside the material simply cannot be established. Accordingly, it is reasonable in most applications in optics, to take  $\mu = 1$ .

The volume charge density  $\rho$  and the current density  $\mathbf{j}(\mathbf{R}, t)$  are the sources of the electromagnetic radiation. In the absence of free charges in the medium or in zones far from the sources, we have  $\mathbf{j} \equiv 0$  and  $\rho \equiv 0$ . (The symbol “ $\equiv$ ” stands for “defined as . . .”) The current density  $\mathbf{j}$  associated with a charge density  $\rho$  moving with a local velocity  $\mathbf{v}$  is

$$\mathbf{j} = \rho \mathbf{v} \quad (2.1.4)$$

Consistency of Eqs. (2.1.1a–d) requires that

$$\operatorname{div} \mathbf{j} + \frac{\partial \rho}{\partial t} = 0 \quad (2.1.5)$$

which is a consequence of charge conservation. This equation, also termed the *equation of continuity*, can be derived as follows. By taking the divergence of both sides of Eq. (2.1.1d), we obtain

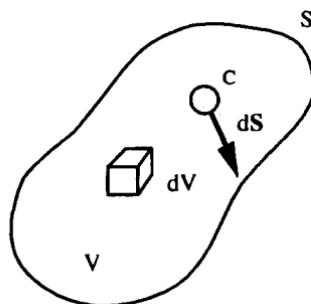
$$\nabla \cdot (\nabla \times \mathbf{H}(\mathbf{R}, t)) = \nabla \cdot \left( \frac{\partial \mathbf{D}}{\partial t}(\mathbf{R}, t) \right) + \nabla \cdot (\mathbf{j}(\mathbf{R}, t)) \quad (2.1.6)$$

Since the left-hand side of Eq. (2.1.6) is zero, we are left with

$$\frac{\partial}{\partial t} \nabla \cdot (\mathbf{D}(\mathbf{R}, t)) + \nabla \cdot (\mathbf{j}(\mathbf{R}, t)) = 0 \quad (2.1.7)$$

which leads to Eq. (2.1.5) by making use of Eq. (2.1.1a). Equation (2.1.5) expresses the fact that the charge is conserved in the neighborhood of any point.

Equivalently, Eqs. (2.1.1a–d) can be cast in integral form. For that purpose, I consider a general volume  $V$  bounded by a simple closed and stationary surface  $S$  (Fig. 2.1.1) and integrate over the volume.



**Figure 2.1.1.** Definition of volume  $V$  bounded by a simple closed surface  $S$  to be used with Maxwell's equations in integral form;  $dS$  is an area element and  $dV$  a volume element. The area  $S$  is bounded by a closed curve  $C$ .

Equation (2.1.1a) can be integrated over a surface  $S$ , as illustrated in Fig. 2.1.1. We find

$$\int_S \mathbf{E} \cdot d\mathbf{S} = \frac{1}{\epsilon\epsilon_0} \int_V \rho \, dV \quad (2.1.8a)$$

which represents the local expression of Gauss' theorem.<sup>4</sup> The outward flux of  $\mathbf{E}$  through a closed surface equals  $1/\epsilon\epsilon_0$  times the enclosed charge. Likewise, the integration of Eq. (2.1.1b) over the area  $S$  yields

$$\int_S \mathbf{B} \cdot d\mathbf{S} = 0 \quad (2.1.8b)$$

The outward flux of  $\mathbf{B}$  through the closed surface  $S$  vanishes: a statement that implies that no free magnetic pole exists. In a similar fashion, Eq. (2.1.1c) can be integrated over a surface  $S$  bounded by  $C$

$$\int_C \mathbf{E} \cdot d\mathbf{l} = - \frac{\partial}{\partial t} \int_S \mathbf{B} \cdot d\mathbf{S} = - \frac{d\Phi}{dt} \quad (2.1.8c)$$

This equation is essentially an integral formulation of the Faraday law of induction. It can be also phrased as follows. Whenever the total net flux through a closed circuit  $C$  varies, there is induced in the circuit a voltage whose magnitude is proportional to the rate of diminution of the flux  $\Phi$  through the

<sup>4</sup>Recall that the divergence of a vector  $\mathbf{a}$  represents physically the outflow of  $\mathbf{a}$  per unit volume through the surface  $S$  enclosing the volume  $V$  and is defined according to

$$\nabla \cdot \mathbf{a} = \lim_{V \rightarrow 0} \frac{1}{V} \int_S \mathbf{a} \cdot d\mathbf{S}$$

circuit. Finally, integrating Eq. (2.1.1d) over  $S$  bounded by a closed curve  $C$ , gives

$$\mu\mu_0 \int_S \left( \mathbf{j} + \epsilon\epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) d\mathbf{S} = \int_C \mathbf{B} \cdot d\mathbf{l} \quad (2.1.8d)$$

One notes that Eq. (2.1.8d) without the displacement current term is Ampere's law, Eq. (2.1.a). The circulation of  $\mathbf{B}$  along a closed curve  $C$  bordering a surface  $S$  equals the flux through  $S$  of the sum of the conduction current and of the displacement current.

To close this section I recall some basic definitions. The flow of electromagnetic energy is given by the Poynting vector

$$\mathbf{P}(\mathbf{R}, t) \equiv \mathbf{E}(\mathbf{R}, t) \times \mathbf{H}(\mathbf{R}, t) \quad (2.1.9)$$

and has units of  $\text{J s}^{-1} \text{m}^{-2}$ . The volume density of energy associated with the electromagnetic field has the value

$$U = \frac{1}{2}(\mathbf{E} \cdot \mathbf{D} + \mathbf{B} \cdot \mathbf{H}) \quad (2.1.10)$$

and has units of  $\text{J m}^{-3}$ . I also define the linear momentum flux associated with the electromagnetic field by the formula

$$\mathbf{G} \equiv \frac{\mathbf{E}(\mathbf{R}, t) \times \mathbf{H}(\mathbf{R}, t)}{c^2} = \frac{\mathbf{P}}{c^2} \quad (2.1.11)$$

## 2.1.2. BOUNDARY CONDITIONS

A unique determination of the field vectors from the set of Eqs. (2.1.1a–c), (2.1.2a–b), (2.1.3) can be obtained in regions of space where both  $\epsilon$  and  $\mu$  are continuous functions of space. In this section I derive relations that the field vectors must satisfy when light is incident on a dielectric medium with different permittivity and permeability, that is, when the physical properties of the medium are discontinuous. The electromagnetic field is uniquely determined within a bounded region  $\Omega$  if one knows the sources of the field inside  $\Omega$  and the values of the electric and magnetic field vectors on the boundaries of the region  $\Omega$ .

At this point we invoke the use of the divergence theorem:

$$\int_V \operatorname{div} \mathbf{a} dV = \int_S \mathbf{a} \cdot d\mathbf{S} \quad (2.1.12)$$

where  $\mathbf{a}$  is an arbitrary field to conclude that Eqs. (2.1.1b) and (2.1.1a) yield

$$\mathbf{n} \cdot (\mathbf{B}_2 - \mathbf{B}_1) = 0 \quad (2.1.13a)$$

$$\mathbf{n} \cdot (\mathbf{D}_2 - \mathbf{D}_1) = \Lambda \quad (2.1.13b)$$

where  $\mathbf{n}$  is the unit vector normal to the surface of discontinuity directed from medium 1 to medium 2 and  $\Lambda$  denotes the surface density of bound charge ( $C m^{-2}$ ). In summary, the normal component of the magnetic induction is continuous. The normal component of the electric displacement changes across the boundary as a result of surface charges.

Likely application of Stokes formula

$$\int_S \mathbf{n} \cdot \text{rot } \mathbf{a} dS = \int_C \mathbf{a} \cdot d\mathbf{l} \quad (2.1.14)$$

where  $\mathbf{a}$  is an arbitrary field that is twice differentiable, to both sides of Eqs. (2.1.1c, d) yields

$$\mathbf{n} \times (\mathbf{E}_2 - \mathbf{E}_1) = 0 \quad (2.1.15a)$$

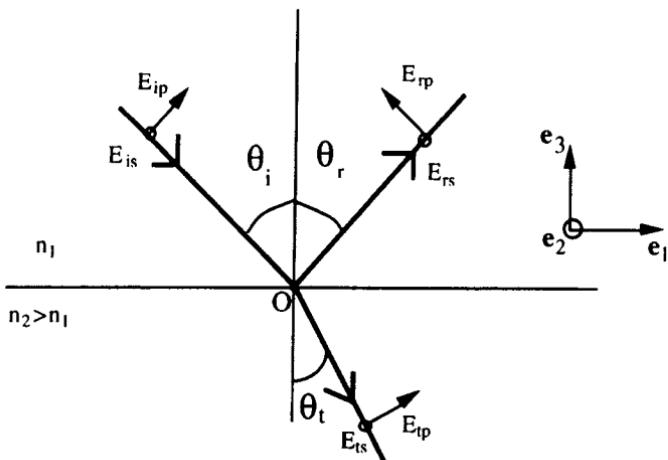
$$\mathbf{n} \times (\mathbf{H}_2 - \mathbf{H}_1) = \kappa \quad (2.1.15b)$$

where  $\kappa$  denotes the surface density of current ( $A m^{-2}$ ). In summary, Eqs. (2.1.15a, b) demonstrate that the tangential components of the electric field vector are continuous across the boundary and that the tangential component of the magnetic vector changes across the boundary as the result of a surface current density.

With these continuity equations as prerequisites, we are now prepared to inquire into the derivation of the Fresnel equations.

### 2.1.3. FRESNEL'S FORMULAS FOR REFLECTION AND TRANSMISSION OF A PLANE WAVE AT A PLANAR BOUNDARY BETWEEN HOMOGENEOUS ISOTROPIC MEDIA

Consider a plane surface or interface at which the physical properties of the medium are discontinuous. The surface is optically smooth, by which is meant that the surface roughness height be small compared with the wavelength of the incident field. When a plane wave falls on this surface at an oblique angle of incidence relative to the normal of this surface, part of the wave energy is reflected and the other part is transmitted (or refracted). In discussing the effects of boundaries, I shall conform to the prevailing convention in the optics literature and find it convenient to adopt the simple geometry of Fig. 2.1.2. This



**Figure 2.1.2.** Reflection and transmission of a plane wave incident on a plane horizontal surface between two homogeneous isotropic media ( $n_1 < n_2$ ). The incident, reflected, and transmitted waves are denoted by the subscripts “i,” “r,” and “t,” respectively. Notation “s” (open symbol) means perpendicular to the page, and notation “p” means in the page. Terms  $\theta_i$ ,  $\theta_r$ , respectively denote the angles of incidence and reflection in the medium of refractive index  $n_1$ , and  $\theta_t$  is the angle of refraction in medium  $n_2 > n_1$ . The incident beam, the normal to the surface, and the reflected beam are all in the same plane.

figure is a schematic depiction of the planar surface separating two semiinfinite homogeneous, isotropic media that are macroscopically characterized by their permittivity and permeability. In the following derivation, I assume that  $\epsilon_1$ ,  $\mu_1$ ,  $\epsilon_2$ , and  $\mu_2$  are real. I also assume that the infinite plane interface between the two dielectrics of indices of refraction  $n_1$  and  $n_2$ , is located perpendicular to the 3-axis in the 1–2 plane, with its direction specified by the unit vector  $e_3$  normal to the surface. The refractive indices  $n_1$  and  $n_2$  are independent on the wave direction. This choice defines a right-handed Cartesian coordinate system. The incident radiation is a plane wave traveling in the 1–3 plane, oriented at an angle with respect to the 3-axis. The incident wave is denoted by the subscript i, the reflected wave is denoted by the subscript r, and the transmitted wave is denoted by the subscript t. I assume that the reflected and transmitted waves are also plane. In general, the reflected and transmitted waves will have different amplitude and phase in the plane of light incidence than in the plane perpendicular to it. It is usual to resolve the electric field vector into components that are parallel (called “p-like,” transverse magnetic wave) and perpendicular (called “s-like” for *senkrecht*, i.e., German for perpendicular, transverse electric wave) to the plane defined by the boundary’s normal and the direction of propagation of the incident wave’s.

According to the Descartes–Snel's law, the incident angle at point  $O$  is equal to the reflection angle where these angles are measured from the normal to the interface at  $O$ . This law can be stated mathematically by

$$\theta_r = \theta_i \quad (2.1.16)$$

and

$$\frac{\sin(\theta_i)}{\sin(\theta_r)} = \frac{n_2}{n_1} = n \quad (2.1.17)$$

where  $n_i = (\epsilon_i \mu_i)^{1/2}$  ( $i = 1, 2$ ) denote the indices of refraction of media 1 and 2, respectively. Here, for future convenience, I introduce the relative index of refraction  $n = (\epsilon_2 \mu_2 / \epsilon_1 \mu_1)^{1/2}$ . Equations (2.1.16) and (2.1.17) determine the relative directions of propagation of the waves. Note that Eqs. (2.1.16) and (2.1.17) depend only on the existence of uniform boundary conditions and not on their precise form. It is also interesting to note that when  $n_2 > n_1$ , which is the case considered in Fig. 2.1.2, there is a real angle  $\theta_r < \theta_i$  for every  $\theta_i$ .

Fresnel's equations relate the amplitude of the reflected and transmitted waves from a plane interface between two linear media,  $E_r$  and  $E_t$ , to the amplitude of the incident wave  $E_i$ . These equations are derived by invoking the boundary-continuity requirements of the tangential components of  $\mathbf{E}$  and  $\mathbf{H}$  across the interface and eliminating  $\mathbf{H}$ . See Refs. 1–7 for a detailed derivation. Here we merely state the result, with reference to the sign convention of Fig. 2.1.2. Hence we have the four relations

$$E_{rs} = \frac{n_1 \cos(\theta_i) - \frac{\mu_1}{\mu_2} n_2 \cos(\theta_r)}{n_1 \cos(\theta_i) + \frac{\mu_1}{\mu_2} n_2 \cos(\theta_r)} E_{is} \quad (2.1.18a)$$

$$E_{rp} = \frac{\frac{\mu_1}{\mu_2} n_2 \cos(\theta_i) - n_1 \cos(\theta_r)}{\frac{\mu_1}{\mu_2} n_2 \cos(\theta_i) + n_1 \cos(\theta_r)} E_{ip} \quad (2.1.18b)$$

$$E_{ts} = \frac{2n_1 \cos(\theta_i)}{n_1 \cos(\theta_i) + \frac{\mu_1}{\mu_2} n_2 \cos(\theta_r)} E_{is} \quad (2.1.18c)$$

$$E_{tp} = \frac{2n_1 \cos(\theta_i)}{\frac{\mu_1}{\mu_2} n_2 \cos(\theta_i) + n_1 \cos(\theta_r)} E_{ip} \quad (2.1.18d)$$

By using the law of refraction [Eq. (2.1.17)] and assuming that both media are

nonmagnetic ( $\mu_1 = \mu_2 = 1$ ), which is frequently encountered in many areas of optics, it is a simple matter to obtain

$$E_{rp} = \frac{\tan(\theta_i - \theta_t)}{\tan(\theta_i + \theta_t)} E_{ip} \quad (2.1.19a)$$

$$E_{rs} = -\frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t)} E_{is} \quad (2.1.19b)$$

$$E_{tp} = \frac{2 \sin(\theta_t) \cos(\theta_i)}{\sin(\theta_i + \theta_t) \cos(\theta_i - \theta_t)} E_{ip} \quad (2.1.19c)$$

$$E_{ts} = \frac{2 \sin(\theta_t) \cos(\theta_i)}{\sin(\theta_i + \theta_t)} E_{is} \quad (2.1.19d)$$

Within these assumptions it turns out that Fresnel's equations contain no explicit reference to the materials characteristics but can be expressed as purely geometric relations. In the following it becomes useful to define the amplitude reflection and transmission coefficients, respectively as  $r_s = E_{rs}/E_{is}$ ,  $r_p = E_{rp}/E_{is}$ ,  $t_s = E_{ts}/E_{is}$ , and  $t_p = E_{tp}/E_{is}$ . Thence the Fresnel equations predict the amplitude changes for s [senkrecht (perpendicular)] and p (parallel) components of the electric field, respectively. We recall that we have implicitly assumed real values for  $n_1$  and  $n_2$ . In the general case for which  $n_1$  and  $n_2$  are complex, Fresnel equations will also predict the phase changes of the complex amplitude coefficients.

Let us now proceed to examine in detail the case of normal incidence ( $\theta_i = \theta_t = 0$ ). It is already evident from Eqs. (2.1.18a–d) that the reflection and transmission coefficients for the s and p forms take the form

$$r_s = \frac{1 - n}{1 + n} \quad (2.1.20a)$$

$$r_p = \frac{n - 1}{1 + n} \quad (2.1.20b)$$

$$t_s = \frac{2}{1 + n} \quad (2.1.20c)$$

$$t_p = \frac{2}{1 + n} \quad (2.1.20d)$$

where again I have set  $n = n_2/n_1 > 1$ , the ratio of refractive indices.

The reflectivities are  $R_s = r_s^2$  and  $R_p = r_p^2$  and the transmissivities are  $T_s = n[\cos(\theta_t)/\cos(\theta_i)]t_s^2$  and  $T_p = n[\cos(\theta_t)/\cos(\theta_i)/\cos(\theta_i)t_p^2]$ . For normal incidence, we see from Eqs. (2.1.20a–d) that  $\lim_{n \rightarrow 1} R_s = R_p = 0$  and  $\lim_{n \rightarrow 1} T_s = T_p = 1$ . The law of conservation of energy imposes that  $R_s + T_s = 1$

and  $R_p + T_p = 1$ . Consider, for instance, an air/crown glass ( $n_2 = 1.5$ ) interface—the intensity of the reflected light is evaluated from Eq. (2.1.20a) and we obtain 4% of the incident light.

If we make use of Descartes-Snel law, the reflection and transmission coefficients can be written for nonnormal incidence as

$$r_s = \frac{\cos(\theta_i) - (n^2 - \sin^2(\theta_i))^{1/2}}{\cos(\theta_i) + (n^2 - \sin^2(\theta_i))^{1/2}} \quad (2.1.21a)$$

$$r_p = \frac{n^2 \cos(\theta_i) - (n^2 - \sin^2(\theta_i))^{1/2}}{n^2 \cos(\theta_i) + (n^2 - \sin^2(\theta_i))^{1/2}} \quad (2.1.21b)$$

$$t_s = \frac{2(\cos(\theta_i))^{1/2}(n^2 - \sin^2(\theta_i))^{1/4}}{\cos(\theta_i) + (n^2 - \sin^2(\theta_i))^{1/2}} \quad (2.1.21c)$$

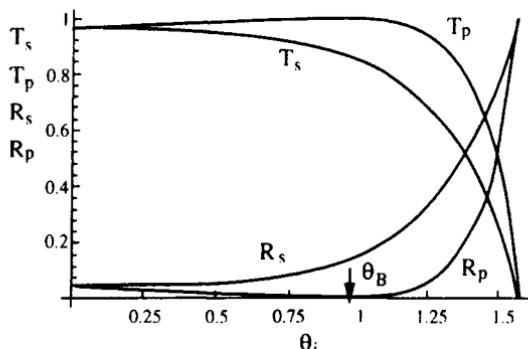
$$t_p = \frac{2n(\cos(\theta_i))^{1/2}(n^2 - \sin^2(\theta_i))^{1/4}}{n^2 \cos(\theta_i) + (n^2 - \sin^2(\theta_i))^{1/2}} \quad (2.1.21d)$$

A word of caution is in order. Here the positive sign of the square root was chosen because the cosine term is positive, real-valued. In the case of interest,  $n > 1$ , the argument of the square root is always positive, real, and less than  $\cos(\theta_i)$ , leading to a negative value of  $r_s$ , that is,  $\pi$  phase shift. The reflection coefficient  $r_p$  goes from a positive value for  $0 < \theta_i < \theta_B$  to a negative value for  $\theta_i > \theta_B$ . When the argument of the square root goes to zero and then becomes negative, the square root becomes imaginary and the negative root is chosen.

The reflectivity and transmittivity will depend on the direction of the incident electric field vector. As an example I show, in Fig. 2.1.3, the angular dependence of the reflectivity and transmittivity for the s and p components as a function of the angle of incidence for an air/crown glass ( $n_2 = 1.5$ ) interface. From inspection of the curves in Fig. 2.1.3, we observe that the intensity of the reflected light remains less than 8% of the incident light for  $\theta_i < \pi/4$  and then it rises to 100% for  $\theta_i = \pi/2$ . Total internal reflection (TIR) occurs when the incident light originates in the dielectric with the larger index of refraction (i.e.,  $n_2$ ) and intersects the interface at an angle of incidence larger than the critical angle  $\theta_c$ :

$$\theta_c \equiv \sin^{-1}(n) \quad (2.1.22)$$

if  $n_2 < n_1$ , where  $n_1$  represents the incidence medium (Fig. 2.1.2). Condition (2.1.22) follows from Eq. (2.1.17) by setting  $\theta_t = \pi/2$ , its maximum value. Then, the reflection coefficients equal one. Thus, over a limited angular range, there is a sharp transition in reflectivity that changes from small value to 1. As an aside, we note that TIR is at the basis for transmitting light into silica fiber consisting of high-refractive-index glass surrounded by a cladding with a lower refractive index.



**Figure 2.1.3.** The reflectivities  $R_s$ ,  $R_p$  and the transmissivities  $T_s$ ,  $T_p$  as functions of the angle of incidence  $\theta_i$  (radians) for an air ( $n_1 = 1$ )/crown glass ( $n_2 = 1.5$ ) interface.

Note also that the intensity of the reflected light for the p-component falls to zero only at  $\theta_i = 56^\circ 39'$ , which is the Brewster angle. At this angle, the intensity of the reflected light for the s-component is about 15%. At Brewster's angle,  $E_{rp} \equiv 0$ , we recognize that

$$\theta_B = \tan^{-1}(n) \quad (2.1.23a)$$

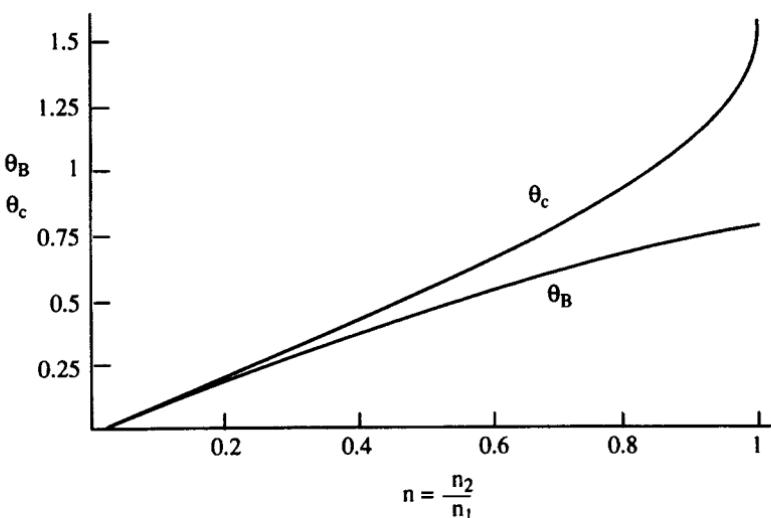
and

$$r_s = \cos(2\theta_B) \quad (2.1.23b)$$

Figure 2.1.4 indicates that  $\theta_c > \theta_B$  and that  $\theta_c \cong \theta_B$  corresponds to small values of  $n = n_2/n_1$ . For media with nearly equal indices of refraction, we have  $\theta_c \cong \pi/2$  and  $\theta_B \cong \pi/4$ .

A result, due to Azzam, indicates that, given two nonmagnetic media of refraction indices  $n_1$  and  $n_2$ , there exists a unique angle of incidence and a unique ratio  $n$  such that, if light is incident on the interface from medium 1, a Brewster angle condition exists (for the p component only), whereas, if the light is incident on the interface from medium 2 (at the same angle), a condition of TIR will exist [8].

In summary, specular reflections can change the state of polarization of an incident beam. That is, the state of polarization of the reflected and transmitted beams is not necessarily the same as that of the incident beam. On the practical side, this behavior leads to a number of applications; for example, intracavity Brewster windows are used in laser technology for polarizing the laser beam. In closing this subsection, it should be mentioned that the problem becomes very bothersome if we are dealing with more general geometries, for example, a nonplane interface [9].



**Figure 2.1.4.** Critical angle  $\theta_c$  (radians) and Brewster angle  $\theta_B$  (radians) as functions of the ratio of refractive indices  $n = n_2/n_1$ . Light is incident on the interface from medium 1.

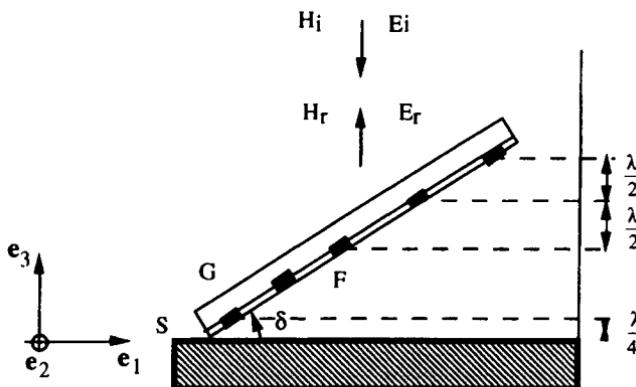
#### 2.1.4. WIENER'S EXPERIMENT

The basic objective of this section is to show that Wiener's experiment provides a clear insight to support the convention that the direction of the electric field vector is chosen to be the polarization direction of the light, a fact that does not seem to be widely discussed in textbooks.

In 1890 Otto Wiener published a remarkable paper (although now largely forgotten) in *Annalen der Physik* (26 years after the work of Maxwell was completed) in which he suggested an ingenious experiment to specify the role of the electric field in optics. Unfortunately, this important experiment has not received the wide recognition it deserves. A brief discussion of this experiment is now given. The reader is referred to Wiener's original paper for details [10].

Wiener's experiment is depicted in Fig. 2.1.5. A beam of monochromatic light of wavelength  $\lambda$  is incident on a plane mirror  $\Sigma$  (silvered on the front surface) in the plane  $x_3 = 0$ . A transparent photographic thin ( $< \lambda/20$ ) film F coated on a glass plate G is placed as illustrated in Fig. 2.1.5 with a small angle  $\delta$  (equal to 4 minutes of arc in the original work of Wiener). One of the remarkable aspects of Wiener's experiment is that the photographic film appeared to be blackened at regular bands.

Let  $E_{i(r)}$  and  $H_{i(r)}$  denote the electric and magnetic fields where the subscripts i and r denote incident and reflected waves. From calculations of the spatial



**Figure 2.1.5.** Schematic illustration of the notations of Wiener's experiment. The  $x_1$  and  $x_2$  axes are chosen to form a right-hand triplet with the  $x_3$  axis and the unit vector  $\mathbf{e}_3$  is along the direction of propagation of the incident and reflected electric fields. The origin of the coordinate system is at the reflecting surface  $\Sigma$ .

dependence of the  $(x_1, x_2)$  components of the total field vector amplitudes, one finds that along the  $x_3$ -axis

$$E_l = E_{li} + E_{lr} \sim \sin\left(\frac{2\pi x_3}{\lambda}\right) \quad (2.1.24a)$$

$$H_l = H_{li} + H_{lr} \sim \cos\left(\frac{2\pi x_3}{\lambda}\right) \quad (2.1.24b)$$

with  $l = x_1, x_2$ . Equations (2.1.24a, b) are periodic functions of  $x_3$ ; the nodes (amplitude minima) of  $\mathbf{E}$  are given by  $x_{3m} = m\lambda/2$  with  $m = 0, 1, 2, \dots$ , while the antinodes (amplitude maxima) are  $x_{3m} = m\lambda/2$  with  $m = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots$ . As can be seen, the nodes of  $\mathbf{E}$  are the antinodes of  $\mathbf{H}$ .

Two points are also noteworthy: (1) the reflecting mirror is a node of  $\mathbf{E}$  and (2) the blackening of the film in Wiener's experiment was found to occur only at the positions of the antinodes of  $\mathbf{E}$ . This is of great importance, for it means that the photochemical action is directly related to the electric and not the magnetic vector. It is interesting to quote Born and Wolf on this subject: "this is to be expected from electron theory since the photographic process is an ionization process in which an electron is removed from an atomic bond of silver halide, and the electromagnetic force on a charged particle at rest is proportional to the electric vector" [6]. Observe that other types of detectors such as fluorescent films and photoemissive films were used in place of the photographic emulsion employed by Wiener; the maximum response was

found in all cases at the *anti-nodes* of the electric field. Most of the experimental techniques are related to the electric field via intensity (observable at optical frequencies, which is a quadratic function of  $E$ ) measurement and in most interactions between light and matter (i.e., the effect of light on a photographic emulsion as in Wiener's experiment), it is the electric vector that is active.

## SECTION 2.2

# Invariance Transformations

Fundamental physical laws are subjected to a variety of symmetry transformations. I now review basic invariance properties of Maxwell's equations. These equations are invariant under gauge transformation, rotations, spatial inversion, and time reversal and are consistent with all experimental facts. We can divide these symmetry transformation to two categories: the dynamical symmetries (e.g., gauge transformation) and the space-time symmetries (e.g., time reversal). A consequence of Noether's theorem is that every continuous symmetry of a system implies a conservation theorem. Dynamical symmetries refer to a particular mathematical transformation of the electromagnetic fields that leaves Maxwell's equations invariant. The conserved quantity corresponding to this symmetry is the electric charge. The other class of symmetries is geometric in nature; the invariance under time translation implies energy conservation, spatial translation implies conservation of momentum, and spatial rotation symmetry implies conservation of angular momentum.

### 2.2.1. GAUGE TRANSFORMATIONS

A fundamental property of Maxwell's equations is that they are invariant under local transformation of potentials

$$\mathbf{E} = -\nabla\varphi - \frac{\partial \mathbf{A}}{\partial t} \quad (2.2.1a)$$

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (2.2.1b)$$

It is important to note that the scalar potential  $\varphi$  and the vector potential  $\mathbf{A}$  are not unique but the different potentials that can be defined all lead to the same physically observable prediction of the theory. This gauge transformation is called *gauge invariance*.

These potentials are coupled by a set of two differential equations that can be obtained by substituting Eqs. (2.2.2a) and (2.2.1b) into Eqs. (2.1.1a) and (2.1.1d). For the fields in vacuum we find

$$\nabla^2\varphi + \frac{\partial}{\partial t}(\nabla \cdot \mathbf{A}) = -\frac{\rho}{\epsilon_0} \quad (2.2.2a)$$

and

$$\nabla^2 \mathbf{A} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} - \nabla \left( \nabla \cdot \mathbf{A} + \epsilon_0 \mu_0 \frac{\partial \varphi}{\partial t} \right) = -\mu_0 \mathbf{j} \quad (2.2.2b)$$

The electric field strength  $\mathbf{E}$ , the magnetic field strength  $\mathbf{B}$ , and Maxwell equations are invariant under the gauge transformations:

$$\varphi(\mathbf{R}) \rightarrow \varphi(\mathbf{R}) - \frac{\partial \Lambda}{\partial t} \quad (2.2.3a)$$

$$\mathbf{A} \rightarrow \mathbf{A} + \mathbf{c} \quad (2.2.3b)$$

for any constant  $c$  and

$$\mathbf{A} \rightarrow \mathbf{A} + \nabla(\Lambda) \quad (2.2.3c)$$

where  $\Lambda$  is an arbitrary function possessing well-defined second-order space and time derivatives. Equations (2.2.3a–b) mean that the definitions of  $\mathbf{A}$  and  $\varphi$  are not unique. The set of Eqs. (2.2.3a–b) forms a group transformation with respect to the function  $\Lambda$ . It is further important to note that the Poynting vector is gauge-independent.

Whereas the scalar potential  $\varphi$  depends on the distribution of charges

$$\varphi(\mathbf{R}, t) = \frac{1}{4\pi\epsilon_0} \int_V \frac{1}{|\mathbf{R}'|} \rho(\mathbf{R}', t) d\mathbf{R}' \quad (2.2.4)$$

the vector potential  $\mathbf{A}$  depends on the distribution of currents

$$\mathbf{A}(\mathbf{R}, t) = \frac{\mu_0}{4\pi} \int_V \frac{1}{|\mathbf{R}'|} \mathbf{j}(\mathbf{R}', t) d\mathbf{R}' \quad (2.2.5)$$

Different conditions on the potentials define specific gauges. The condition defining the *Coulomb gauge* is  $\nabla \cdot \mathbf{A} = 0$ . Then Eqs. (2.2.1) and (2.2.2) read

$$\nabla^2 \varphi = -\frac{\rho}{\epsilon_0} \quad (2.2.6a)$$

known as *Poisson's equation*, and

$$\nabla^2 \mathbf{A} - \epsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} - \epsilon_0 \mu_0 \nabla \left( \frac{\partial \varphi}{\partial t} \right) = -\mu_0 \mathbf{j} \quad (2.2.6b)$$

For the *Lorentz gauge* we impose the condition  $\nabla \cdot \mathbf{A} + \epsilon_0 \mu_0 (\partial \varphi / \partial t) = 0$ . Now

Eqs. (2.2.1) and (2.2.2) become uncoupled and read

$$\nabla^2 \varphi - \varepsilon_0 \mu_0 \frac{\partial^2 \varphi}{\partial t^2} = -\frac{\rho}{\varepsilon_0} \quad (2.2.7a)$$

and

$$\nabla^2 \mathbf{A} - \varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} = -\mu_0 \mathbf{j} \quad (2.2.7b)$$

A feature to be particularly emphasized is that in free space, that is, in a region devoid of sources, both the Coulomb gauge and the Lorentz gauge are satisfied and we can represent the electromagnetic field using only a vector potential  $\mathbf{A}$ .

Given a source of volume  $V$ , the solution of Eq. (2.2.6) or (2.2.7) at point  $P$ , characterized by position vector  $\mathbf{R}$  depends on the source density at the point  $M$ , characterized by position vector  $\mathbf{R}'$ , as displayed in Fig. 2.2.1.

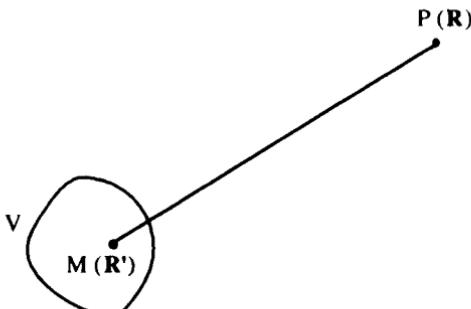
We get

$$\varphi(\mathbf{R}, t) = \frac{1}{4\pi\varepsilon_0} \int_V \frac{1}{|\mathbf{R} - \mathbf{R}'|} \rho(\mathbf{R}', \tau) d\mathbf{R}' \quad (2.2.8)$$

and

$$\mathbf{A}(\mathbf{R}, t) = \frac{\mu_0}{4\pi} \int_V \frac{1}{|\mathbf{R} - \mathbf{R}'|} \mathbf{j}(\mathbf{R}', \tau) d\mathbf{R}' \quad (2.2.9)$$

where  $\tau = t - R\sqrt{\varepsilon_0\mu_0}$ . Equations (2.2.8) and (2.2.9) are known as the *retarded potential solutions*, that is,  $R\sqrt{\varepsilon_0\mu_0}$  is the time needed for the wave to travel from  $\mathbf{R}'$  to  $\mathbf{R}$ . Then  $\varphi$  (resp.  $\mathbf{A}$ ) depends on the charge density (resp. current



**Figure 2.2.1.** Schematic illustration of the notation related to the retarded potential solutions.

density) everywhere in space, and charge densities (resp. current densities) are evaluated at the appropriate retarded times. For the harmonic case, namely,  $\rho(\mathbf{R}, t) = \rho(\mathbf{R}) \exp(-2i\pi v t)$ ,  $\mathbf{j}(\mathbf{R}, t) = \mathbf{j}(\mathbf{R}) \exp(-2i\pi v t)$  we have

$$\varphi(\mathbf{R}, t) = \frac{1}{4\pi\epsilon_0} \int_V \frac{\exp(-ik|\mathbf{R} - \mathbf{R}'|)}{|\mathbf{R} - \mathbf{R}'|} \rho(\mathbf{R}') d\mathbf{R}' \quad (2.2.10)$$

$$\mathbf{A}(\mathbf{R}, t) = \frac{\mu_0}{4\pi} \int_V \frac{\exp(-ik|\mathbf{R} - \mathbf{R}'|)}{|\mathbf{R} - \mathbf{R}'|} \mathbf{j}(\mathbf{R}') d\mathbf{R}' \quad (2.2.11)$$

where  $k \equiv 2\pi v \sqrt{\epsilon_0 \mu_0}$  and  $v$  is the frequency.

## 2.2.2. TIME INVERSION

The specific role of the electric field in optics is also based on the time-reversal symmetry ( $T$  symmetry). Fundamental physical laws are subject to a variety of symmetry transformations. Their symmetry properties impose restrictions on the possible properties of matter and fields. The basic laws of classic physics are invariant to the sense of direction of time (or, equivalently, under reversal of the direction of motion).

The reversibility in classic point mechanics can be expressed in the following way. Let us call two states of a mechanical system mutually reversed states, assuming that particles have the same position and the opposite velocities. Then the reversibility of mechanics means that if a mechanical system that was in the state  $S$  at time  $t$  finds itself in the state  $S'$  at time  $t'$ , then the fundamental laws allow for another solution representing the similar system that was in the reversed state of  $S'$  at time  $t$  and that finds itself in the reversed state  $S$  at time  $t'$ . This means that, under the time-reversal transformation  $t \rightarrow -t$ , the related form of the equations is the same as before. To extend the concept of reversibility to electrodynamics, I need to add to the definition of reversed states the condition that the electric field has the same value and the magnetic field has the same absolute value but opposite sign. Since the equations of motion of classic mechanics are differential equations involving second-order temporal derivatives, they are invariant under the  $T$ -symmetry transformation  $t \rightarrow -t$  while odd functions of  $t$  such as the velocity  $\mathbf{v}$  change of sign. As a result, a main difference arises between static and dynamic quantities. Thus at point  $\mathbf{R}$ , the charge density  $\rho(\mathbf{R}, t)$  and the current density  $\mathbf{j}(\mathbf{R}, t)$  behave differently under  $T$  symmetry:  $\rho(\mathbf{R}, t) \rightarrow \rho(\mathbf{R}, -t)$ ;  $\mathbf{j}(\mathbf{R}, t) \rightarrow -\mathbf{j}(\mathbf{R}, -t)$ .

Now let us consider the equation of motion of a point charge  $q$  of mass  $m$  moving in vacuum, with speed  $\mathbf{v}$ , under the influence of the Lorentz force in an electric field  $\mathbf{E}(\mathbf{R}, t)$  and a magnetic field  $\mathbf{H}(\mathbf{R}, t)$ . This equation reads for nonrelativistic speeds ( $|\mathbf{v}|/c \ll 1$ ) as:

$$m \frac{d^2 \mathbf{R}}{dt^2} = q \mathbf{E}(\mathbf{R}, t) + \mu_0 q \mathbf{v} \times \mathbf{H}(\mathbf{R}, t) \quad (2.2.12)$$

It is an experimental fact that electric charge is invariant under Lorentz transformations and rotations. If it is assumed that charge and charge density are scalar under the  $T$  symmetry, then Eq. (2.2.12) will be invariant under the  $T$  symmetry if and only if  $\mathbf{E}(\mathbf{R}, t) \rightarrow \mathbf{E}(\mathbf{R}, -t)$  and  $\mathbf{H}(\mathbf{R}, t) \rightarrow -\mathbf{H}(\mathbf{R}, -t)$ . This conclusion is, of course, to be expected from Maxwell electromagnetic field theory. The electric field is produced by a distribution of static charges unaffected by the  $T$  symmetry while the magnetic field originates from a distribution of moving electrical charges (current density) whose sign is reversed by the  $T$  symmetry. The electric and magnetic fields do not appear symmetrically in the Maxwell and Lorentz equations. This important point was emphasized by Rosen [11] as follows:

Whereas there exist in nature electric charges which serve as sources of the electric fields and detectors of it, when in motion serve as sources of the magnetic field and detectors of it also, there do not exist, at least as far as is presently known, magnetic charges (magnetic monopoles) which at rest would produce and detect magnetic fields and in motion would also product and detect electric field. If such magnetic charges are ever discovered, then the electric and magnetic fields would be completely symmetric.

If one assumes that  $\mathbf{E}$  and  $\mathbf{H}$  are plane harmonic waves connected by way of the Maxwell equations, one can show from Eq. (2.2.12) that the ratio of the magnitude of the magnetic force to that of the electric force on the point charge  $q$  cannot exceed  $|v|/c$ . Since, in the range of optical frequencies, electrons have only small velocities, the magnetic effects are usually unimportant and can quite correctly be ignored. Thus the electric field of light is what exerts most of the forces on charges in matter. In this sense, the electric field is responsible for not only changes of polarization state on interaction of light with matter but also photography, vision, the detection of light, and much else, namely, Wiener's experiment.

### 2.2.3. SPATIAL INVERSION

The requirement of symmetry under spatial inversion can be derived from Eqs. (2.1.1)–(2.1.4). Under spatial inversion  $\mathbf{r} \rightarrow -\mathbf{r}$  and  $\mathbf{v} \rightarrow -\mathbf{v}$ , then the set of Maxwell's equations will be invariant if and only if  $\mathbf{E}(\mathbf{R}, t) \rightarrow -\mathbf{E}(-\mathbf{R}, t)$ , and  $\mathbf{H}(\mathbf{R}, t) \rightarrow \mathbf{H}(-\mathbf{R}, t)$ .

### 2.2.4. DUALITY TRANSFORMATIONS

Again it is helpful to consider the vacuum Maxwell equations for the electric and magnetic fields. Here the basic point I wish to stress is that if the following transformations of the fields are made

$$\mathbf{E} \rightarrow -\mu_0 \mathbf{H}, \quad \mathbf{H} \rightarrow \epsilon_0 \mathbf{E} \quad (2.2.13)$$

or

$$\begin{aligned}\mathbf{E} &\rightarrow \mathbf{E} \cos(\theta) + \mathbf{H} \sin(\theta) \\ \mathbf{H} &\rightarrow \mathbf{H} \cos(\theta) - \mathbf{E} \sin(\theta)\end{aligned}\tag{2.2.14}$$

then Maxwell's equations are also satisfied. Note that Eq. (2.2.14) is also satisfied for the paired fields ( $\mathbf{B}, \mathbf{D}$ ) instead of ( $\mathbf{E}, \mathbf{H}$ ). This symmetry still holds in the presence of charges and currents if one adds both electric and magnetic charges and currents.

## SECTION 2.3

# Monochromatic Plane-Wave

The purpose of this section is to find plane-wave solutions to Maxwell's Eq. (2.1.1a-d) in source-free space and to review their properties.

### 2.3.1. HARMONIC PLANE-WAVE SOLUTION OF THE WAVE EQUATION

I consider regions far from the sources where both  $\rho \cong 0$  and  $\mathbf{j} \cong 0$ , which is often encountered in many areas in optics. In a homogeneous and isotropic medium with no free charges ( $\rho = 0$ ) and no current density present ( $\mathbf{j} = 0$ ), it follows from Eq. (2.2.6b) or (2.2.7b) that the potential vector  $\mathbf{A}(\mathbf{R}, t)$ , at a point specified by position vector  $\mathbf{R}$  and at time  $t$ , satisfies the Helmholtz equation

$$\nabla^2 \mathbf{A} - \epsilon\epsilon_0\mu\mu_0 \frac{\partial^2 \mathbf{A}}{\partial t^2} = 0 \quad (2.3.1)$$

where  $\mathbf{A}$  also satisfies  $\nabla \cdot \mathbf{A} = 0$ .

Equation (2.3.1) is the wave equation that describes light propagation. Plane and spherical waves are general solutions of Eq. (2.3.1); that is, their derivations do not require boundary conditions. We see from Eqs. (2.1.1a, b) that

$$\mathbf{E} = - \frac{\partial \mathbf{A}}{\partial t} \quad (2.3.2)$$

and

$$\mathbf{B} = \nabla \times \mathbf{A} \quad (2.3.3)$$

A wave equation identical to Eq. (2.3.1) holds for the electric field  $\mathbf{E}$  and the magnetic field  $\mathbf{H}$ . It is worth noting that not every electric field that is a solution of the wave equation is a solution of Maxwell's equations. One must also require that the field must satisfy  $\nabla \cdot \mathbf{E} = 0$  in order to be a solution of Maxwell's equations.

Although no optical source produces strictly monochromatic light, I attempt now to study in detail the harmonic solution of Maxwell's equations

for at least two reasons: (1) from the linearity of these equations, it follows that a harmonic excitation at a definite frequency  $v$  produces a harmonic response and (2) Fourier analysis shows that any transient function can be represented by a superposition of harmonic components. When the fields are harmonic, it is convenient to represent all variables in complex form, with the understanding that the real parts of the complex quantities represent the physical values of the fields. I will denote the real part of the complex vector  $\mathbf{E}$  by

$$\mathbf{E}^{(r)}(\mathbf{R}, t) = \operatorname{Re}(\mathbf{E}(\mathbf{R}, t)) \quad (2.3.4)$$

Assume that the electric field  $\mathbf{E}(\mathbf{R}, t) = \mathbf{E}(\mathbf{R}) \exp(-2i\pi vt)$  and the magnetic field  $\mathbf{H}(\mathbf{R}, t) = \mathbf{H}(\mathbf{R}) \exp(-2i\pi vt)$  are solutions of Maxwell's equations with  $i^2 = -1$ . We now exploit the fact that we can factor out the time dependence in Eqs. (2.1.1a-d). Hence the Maxwell equations reduce to differential equations in the spatial variable only. By substituting these complex representations into Maxwell's equations we get the following homogeneous equations

$$\nabla \cdot (\epsilon \epsilon_0 \mathbf{E}) = 0 \quad (2.3.5a)$$

$$\nabla \cdot (\mu \mu_0 \mathbf{H}) = 0 \quad (2.3.5b)$$

$$\nabla \times \mathbf{E} = 2i\pi v \mu \mu_0 \mathbf{H} \quad (2.3.5c)$$

$$\nabla \times \mathbf{H} = -2i\pi v \epsilon \epsilon_0 \mathbf{E} \quad (2.3.5d)$$

It is a simple matter to combine these equations to obtain the Helmholtz equation for the electric field in a spatially uniform medium

$$\nabla^2 \mathbf{E} + 4\pi^2 v^2 \epsilon \epsilon_0 \mu \mu_0 \mathbf{E} = 0 \quad (2.3.6)$$

Assume now that the direction of propagation of the lightwave is along  $\mathbf{e}_3$ , then the electric field vector lies in the plane  $(\mathbf{e}_1, \mathbf{e}_2)$ . If we further require a spatial dependence for  $\mathbf{E}$  of the form  $\exp(i\mathbf{q} \cdot \mathbf{R})$ , we have the following complex representation of a monochromatic wave

$$\mathbf{E}_j(\mathbf{R}, t) = a_j \exp(\pm i(\mathbf{q} \cdot \mathbf{R} - 2\pi vt + \theta_j)) = a_j \exp(\pm i(q\mathbf{e}_3 \cdot \mathbf{R} - 2\pi vt + \theta_j)) \quad (2.3.7)$$

where the  $a_j$  values  $j = 1, 2$  are complex amplitudes that are constant in space and time, the quantity  $\mathbf{q} \cdot \mathbf{R} - 2\pi vt + \theta_j$  is the phase of the wave, and  $\mathbf{q}$  denotes the wave vector. The plane containing the electric field vector  $\mathbf{E}$  and to which the wave vector  $\mathbf{q}$  is perpendicular is known as the plane of vibration. At fixed  $t$ , the plane wave solution is not generally periodic with respect to space, that is, damped wave, whereas at any fixed position Eq. (2.3.7), is periodic in time with a period  $1/v$ . Consider the case in which the wave vector  $\mathbf{q}$  is real. For this type of wave, the electric field  $\mathbf{E}(\mathbf{R}, t)$  takes the same value everywhere on a surface of constant phase  $\mathbf{q} \cdot \mathbf{R} - 2\pi vt + \theta_j = \text{constant}$ , called a wavefront. At

a given instant of time, the wavefront is a plane normal to  $\mathbf{q}$  and is also a plane of constant amplitude. The plane wave is said to be homogeneous. Let us now consider a time harmonic plane wave with a complex wave vector  $\mathbf{q} = \text{Re}(\mathbf{q}) + i\text{Im}(\mathbf{q})$ . In contrast with the previous case, the planes of constant amplitude (i.e.,  $\text{Im}(\mathbf{q}) \cdot \mathbf{R} = \text{constant}$ ) of the wave do not, in general, coincide with the planes of constant phase (i.e.,  $\text{Re}(\mathbf{q}) \cdot \mathbf{R} = \text{constant}$ ). A plane wave of this type is called an inhomogeneous plane wave [6].

In the following we restrict our attention to the case of homogeneous waves propagating in a material medium characterized by real  $\epsilon$  and  $\mu$ . From Maxwell's equations, we can also write the vector fields of the plane wave in the form

$$\mathbf{E}(\mathbf{R}, t) = a_1 \mathbf{e}_1 \exp(\pm i(q\mathbf{e}_3 \cdot \mathbf{R} - 2\pi v t + \theta_1)) + a_2 \mathbf{e}_2 \exp(\pm i(q\mathbf{e}_3 \cdot \mathbf{R} - 2\pi v t + \theta_2)) \quad (2.3.8)$$

and

$$\begin{aligned} \mathbf{H}(\mathbf{R}, t) &= \frac{\mathbf{q} \times \mathbf{E}(\mathbf{R}, t)}{2\pi v \mu \mu_0} = \left( \frac{\mathbf{q} \times \mathbf{e}_1}{2\pi v \mu \mu_0} \right) a_1 \exp(\pm i(q\mathbf{e}_3 \cdot \mathbf{R} - 2\pi v t + \theta_1)) \\ &\quad + \left( \frac{\mathbf{q} \times \mathbf{e}_2}{2\pi v \mu \mu_0} \right) a_2 \exp(\pm i(q\mathbf{e}_3 \cdot \mathbf{R} - 2\pi v t + \theta_2)) \end{aligned} \quad (2.3.9)$$

The electric and magnetic fields are transverse (i.e.,  $\mathbf{E} \perp \mathbf{q}$ ,  $\mathbf{H} \perp \mathbf{q}$ ): orthogonal to each other (i.e.,  $\mathbf{E} \perp \mathbf{H}$ ); in phase and their magnitudes are related by  $|\mathbf{E}| = |\mathbf{H}|(\mu \mu_0 / \epsilon \epsilon_0)^{1/2}$  (in vacuum  $|\mathbf{E}/\mathbf{H}| = (\mu_0/\epsilon_0)^{1/2} = 377 \Omega$ ). In other words,  $\mathbf{q}$ ,  $\mathbf{E}$ , and  $\mathbf{H}$  define a right-handed coordinate system: the direction of the electric field vector is parametrized by unit vectors  $\mathbf{e}_1$  and  $\mathbf{e}_2$ . I remark, parenthetically, that in a general anisotropic medium, only the field vectors  $\mathbf{D}$  and  $\mathbf{B}$  are orthogonal to the direction of propagation.

A point to bear in mind is that the plus or minus sign in the exponential indicates that waves can propagate in the two directions. We then have two choices for assigning the time-dependent part of the exponential. Either choice of sign convention yields the correct physics; however, the time dependence that is chosen imposes certain constraints (i.e., a wave propagating in an absorbing medium characterized by a complex index of refraction is always attenuated); hence, if a negative time exponent is chosen, the imaginary part of the complex refractive index must be positive. Here the sign convention used is that waves have a negative imaginary time exponent. This sign convention is discussed in detail in Ref. 12.

By making the substitution of Eq. (2.3.8) into Eq. (2.3.6), we see at once the following relation between the modulus of the complex wave vector  $\mathbf{q}$  and the amplitude

$$|\mathbf{q}| = 2\pi v(\epsilon \epsilon_0 \mu \mu_0)^{1/2} \quad (2.3.10)$$

This equation is known as the *dispersion relation* for the magnitude of  $\mathbf{q}$ . The constant electric field plane at  $\mathbf{R}$  is defined by  $\mathbf{q} \cdot \mathbf{R}$ . The planes with equal phases are spaced apart by the wavelength  $\lambda \equiv v/\nu$ . The phase velocity is the velocity of the surface of constant phase; thus, as time progresses, the planes normal to the wave vector  $\mathbf{q}$  move with velocity

$$v = 2\pi \frac{\nu}{|\mathbf{q}|} \quad (2.3.11)$$

Moreover, in a source-free space, there is no dispersion of the waves. Therefore the group velocity  $2\pi(d\nu/d|\mathbf{q}|)$  is equal to the phase velocity  $v$ . From Eq. (2.3.10) and (2.3.11) we obtain the expression of the phase velocity:

$$v = \frac{1}{\sqrt{\epsilon\epsilon_0\mu\mu_0}} \quad (2.3.12)$$

In vacuum, the velocity of electromagnetic radiation has the value

$$c = \frac{1}{\sqrt{\mu_0\epsilon_0}} = 2.9979 \cdot 10^8 \text{ m s}^{-1} \quad (2.3.13)$$

In an infinite plane wave the  $\mathbf{E}$  and  $\mathbf{H}$  fields are everywhere perpendicular to the wave vector and the energy flow is everywhere parallel to the wave vector. For completeness, I indicate that the index of refraction of the material medium reduces to

$$n = \frac{c}{v} = (\epsilon\mu)^{1/2} \quad (2.3.14)$$

Since the frequencies we are interested in are high ( $\sim 10^{15}$  Hz), it is usual to consider time-averaged values rather than instantaneous values of relevant physical quantities. This is because detectors cannot respond to instantaneous values of the Poynting vector. It is important to appreciate that if we want to compute the time average of the Poynting vector for the plane-wave solution of Maxwell's equations, we must take the real parts of the complex instantaneous fields before computing the vector product in Eq. (2.1.9). This follows from the fact that the real part of the product of two complex numbers is not the product (in general) of their real parts. The time-averaged Poynting vector over a number of complete oscillations is thus defined by

$$\langle \mathbf{P}(\mathbf{R}, t) \rangle = \langle \mathbf{E}^{(r)}(\mathbf{R}, t) \times \mathbf{H}^{(r)}(\mathbf{R}, t) \rangle = \frac{1}{2}\epsilon_0 c \left(\frac{\epsilon}{\mu}\right)^{1/2} \langle |\mathbf{E}|^2 \rangle \frac{\mathbf{q}}{|\mathbf{q}|} \quad (2.3.15)$$

where I made use of the following theorem: The time average (denoted by angular brackets), over a time that is large compared with the inverse

frequency of the product of two harmonic time-dependent functions  $a$  and  $b$ , of the same frequency  $\nu$ , can be written as

$$\begin{aligned}\langle a(t) \times b(t) \rangle &= \frac{1}{T} \int_0^T \frac{1}{4} \left[ a \exp(-2i\pi\nu t) + a^* \exp(2i\pi\nu t) \right] \\ &\quad \times [b \exp(-2i\pi\nu t) + b^* \exp(2i\pi\nu t)] dt \quad (2.3.16) \\ &= \frac{1}{2} \operatorname{Re}(a \times b^*) = \frac{1}{2} \operatorname{Re}(a^* \times b)\end{aligned}$$

where I have set  $T = 1/\nu$  the period of oscillation. [The asterisk (\*) denotes complex conjugate.]

Another way of introducing the Poynting vector and the volume density of energy is in terms of complex notation. We have

$$\mathbf{P}(\mathbf{R}, t) = \frac{1}{2} \operatorname{Re}(\mathbf{E}(\mathbf{R}, t) \times \mathbf{H}^*(\mathbf{R}, t)) \quad (2.3.17)$$

and

$$U = \frac{1}{4} \operatorname{Re}(\mathbf{E} \cdot \mathbf{D}^* + \mathbf{B} \cdot \mathbf{H}^*) \quad (2.3.18)$$

For a uniform plane wave propagating in the positive direction of the 3-axis,  $\mathbf{E}$  is independent of  $x_1$  and  $x_2$ ,  $\nabla \cdot \mathbf{E} = \partial \mathbf{E}_3 / \partial x_3 = 0$ , and  $E_3$  cannot be a function of  $x_3$ ; we shall set  $E_3 = 0$ . When  $x_3 = 0$ , we may write the electric field vector of the plane wave in the form

$$E_1^{(r)}(t) = a_1 \cos(\theta_1 - 2\pi\nu t) \quad (2.3.19a)$$

$$E_2^{(r)}(t) = a_2 \cos(\theta_2 - 2\pi\nu t) \quad (2.3.19b)$$

where  $a_1$ ,  $a_2$ ,  $\theta_1$  modulo  $(2\pi)$ , and  $\theta_2$  modulo  $(2\pi)$  represent the constant amplitudes and phases, respectively, and are determined by the boundary conditions.

In closing this subsection I note that in a dispersion-free medium the time-averaged electric and magnetic energy densities are equal:  $\epsilon\epsilon_0|\mathbf{E}^2| = \mu\mu_0|\mathbf{H}^2|$ .

### 2.3.2. RECIPROCITY PRINCIPLE

To derive the reciprocity principle, I consider a region devoid of sources where the medium is linear and isotropic. Let  $\mathbf{E}_1$ ,  $\mathbf{H}_1$  and  $\mathbf{E}_2$ ,  $\mathbf{H}_2$  be harmonic field vectors of the same frequency for two independent solutions of Maxwell's equations [Eqs. (2.1.9a, b)]. They satisfy

$$\nabla \times \mathbf{H}_j = -2i\pi\nu\epsilon\epsilon_0\mathbf{E}_j \quad (2.3.20a)$$

and

$$\nabla \times \mathbf{E}_j = 2i\pi\nu\mu\mu_0\mathbf{H}_j \quad (2.3.20b)$$

We first set  $j = 1$  and dot-multiply Eq. (2.3.20a) by  $\mathbf{E}_2$ , then Eq. (2.3.20b) by  $\mathbf{H}_2$  and add. We have

$$\nabla \times \mathbf{H}_1 \cdot \mathbf{E}_2 + \nabla \times \mathbf{E}_1 \cdot \mathbf{H}_2 = -2i\pi\nu(\epsilon\epsilon_0\mathbf{E}_1 \cdot \mathbf{E}_2 - \mu\mu_0\mathbf{H}_1 \cdot \mathbf{H}_2) \quad (2.3.21)$$

Interchanging the subscript  $1 \leftrightarrow 2$  in Eq. (2.3.21), we find

$$\nabla \times \mathbf{H}_2 \cdot \mathbf{E}_1 + \nabla \times \mathbf{E}_2 \cdot \mathbf{H}_1 = -2i\pi\nu(\epsilon\epsilon_0\mathbf{E}_1 \cdot \mathbf{E}_2 - \mu\mu_0\mathbf{H}_1 \cdot \mathbf{H}_2) \quad (2.3.22)$$

By subtracting Eqs. (2.3.22) and (2.3.21) and using a standard vector identity for the divergence of a vector product on the left-hand side, we see that

$$\nabla \cdot (\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1) = 0 \quad (2.3.23)$$

This result is also known as *Lorentz's lemma*. Next, integrating over a volume  $V$  bounded by a stationary surface  $S$  and using Gauss' theorem, we obtain

$$\int_S \mathbf{E}_1 \times \mathbf{H}_2 \cdot d\mathbf{S} = \int_S \mathbf{E}_2 \times \mathbf{H}_1 \cdot d\mathbf{S} \quad (2.3.24)$$

This equation is the integral statement of the reciprocity theorem.

### 2.3.3. POYNTING'S THEOREM

Dot multiplying both sides of Eq. (2.1.1d) by  $\mathbf{E}$ , we can calculate the rate of work done by the vector fields per unit volume, defined as

$$\mathbf{j} \cdot \mathbf{E} = \mathbf{E} \cdot (\nabla \times \mathbf{H}) - \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \quad (2.3.25)$$

which reduces to

$$\mathbf{j} \cdot \mathbf{E} = -\nabla \cdot (\mathbf{E} \times \mathbf{H}) - \mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} - \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \quad (2.3.26)$$

if we make use of the formula

$$\nabla \cdot (\mathbf{a} \times \mathbf{b}) = \mathbf{b} \cdot \nabla \times \mathbf{a} - \mathbf{a} \cdot \nabla \times \mathbf{b} \quad (2.3.27)$$

where  $\mathbf{a}$  and  $\mathbf{b}$  are two arbitrary vectors. By substitution of Eq. (2.3.27) into

Eq. (2.3.26), we obtain the modified equation of continuity

$$\operatorname{div} \mathbf{P} + \frac{\partial U}{\partial t} = -\mathbf{j} \cdot \mathbf{E} \quad (2.3.28)$$

This equation is a statement of the Poynting theorem and has a definite physical meaning in terms of energy conservation: This energy-balance equation states that, while the total energy is conserved, the electromagnetic energy may not be conserved.

I have stated the Poynting theorem in differential form; it can also be cast in integral form. Integrating both sides of Eq. (2.3.26) over a volume  $V$  (with reference to Fig. 2.1), we have

$$\int_S (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{S} + \frac{\partial}{\partial t} \int_V U dV = - \int_V \mathbf{j} \cdot \mathbf{E} dV \quad (2.3.29)$$

The flux of the Poynting vector through the closed surface  $S$  bounding the volume  $V$  is equal to the sum of the time-rate exchange of the energy within the volume  $V$  and of the power associated with the current density  $\mathbf{j}$ .

In the special case of an electromagnetic plane wave propagating in free space, we have, from Eq. (2.1.10)

$$\int_S (\mathbf{E} \times \mathbf{H}) \cdot d\mathbf{S} + \frac{\partial}{\partial t} \int_V \left( \epsilon_0 \frac{E^2}{2} + \mu_0 \frac{H^2}{2} \right) dV = 0 \quad (2.3.30)$$

### 2.3.4. NORMAL MODES REPRESENTATION OF ELECTROMAGNETIC FIELD

For many applications, it is more convenient to deal with a discrete set of variables than the whole continuum at once. For that purpose I shall describe the field localized inside a certain volume of space that we can expand in terms of a discrete set of orthogonal mode functions. From the very property of linearity of Maxwell's equations, each solution of these equations may be written as a superposition of plane waves. The concept of normal mode arises quite naturally in the consideration of a radiation field in a source-free cavity. For that purpose we take the solution of Eq. (2.3.1) by representing the vector potential as a discrete sum of mode functions

$$\mathbf{A}(\mathbf{R}, t) = (\epsilon_0)^{-1/2} \sum_q e_q(t) \mathbf{u}_q(\mathbf{R}) \quad (2.3.31)$$

which are defined according the following wave equations, for each mode  $q$ :

$$\nabla^2 \mathbf{u}_q(\mathbf{R}) = -4\pi^2 \frac{v_q^2}{c^2} \mathbf{u}_q(\mathbf{R}) \quad (2.3.32a)$$

The coefficient  $(\epsilon_0)^{-(1/2)}$  is used for normalization purposes. The amplitude of each mode satisfies

$$\frac{d^2 e_q}{dt^2} + 4\pi^2 v_q^2 e_q = 0 \quad (2.3.32b)$$

Equations (2.3.32a, b) define a set of frequencies  $\{v_q\}$ . The choice of boundary condition is to a considerable degree at one's disposal and can be determined by mathematical convenience. It is convenient to consider a cubic cavity of side  $L$ , which is very large compared with the dimensions of the region of interest. If we impose boundary conditions within the cubic cavity, the vector potential has an infinite discrete and complete set of normal-mode solutions with the orthonormality condition

$$\int_{\text{cavity}} \mathbf{u}_q(\mathbf{R}) \cdot \mathbf{u}_{q'}(\mathbf{R}) dV = \delta_{qq'} \quad (2.3.33)$$

where the integration is carried out over the volume of the cavity. From the gauge invariance we have also  $\nabla \cdot (\mathbf{u}_q(\mathbf{R})) = 0$  everywhere in the cavity. The basis functions  $\mathbf{u}_q(\mathbf{R})$  are  $\exp(i\mathbf{q} \cdot \mathbf{R})$  for the cubical cavity.

The electromagnetic energy inside the cavity is from Eq. (2.1.10),

$$\begin{aligned} H &= \int_{\text{cavity}} U dV = \frac{1}{2} \int_{\text{cavity}} (\epsilon_0 \mathbf{E}^2 + \mu_0 \mathbf{H}^2) dV \\ &= \frac{1}{2} \int_{\text{cavity}} \left( \epsilon_0 \left( \frac{\partial \mathbf{A}}{\partial t} \right)^2 + \mu_0 (\nabla \times \mathbf{A})^2 \right) dV \end{aligned} \quad (2.3.34)$$

By substituting Eq. (2.3.31) into Eq. (2.3.34), we obtain

$$H = \frac{1}{2} \dot{e}_q \dot{e}_{q'} \int_{\text{cavity}} \mathbf{u}_q \cdot \mathbf{u}_{q'} dV + \frac{c^2}{2} \sum_{k,l} e_k e_{k'} \int_{\text{cavity}} (\nabla \times \mathbf{u}_q) \cdot (\nabla \times \mathbf{u}_{q'}) dV \quad (2.3.35)$$

Derivatives with respect to time are designated with overdots. By making use of a standard vector identity,<sup>5</sup> Gauss' theorem, and a suitable set of boundary equations, we see that Eq. (2.3.35) reduces to

$$H = \frac{1}{2} \sum_q (\dot{e}_q^2 + 4\pi^2 v_q^2 e_q^2) = \sum_q H_q \quad (2.3.36)$$

We can readily check that  $H_q$  represents the energy of a harmonic oscillator of frequency  $v_q$ . With respect to this discrete basis, we may specify completely the

<sup>5</sup>The vector identity reads  $(\nabla \times \mathbf{u}_q) \cdot (\nabla \times \mathbf{u}_{q'}) = \mathbf{u}_q \cdot \nabla \times (\nabla \times \mathbf{u}_{q'}) + \nabla \cdot (\mathbf{u}_{q'} \times (\nabla \times \mathbf{u}_q))$ . A reader wishing to pursue the details of the calculation is advised to consult Louisell [13].

electric and the magnetic fields by the set of amplitudes  $\{e_q\}$

$$\mathbf{E}(\mathbf{R}, t) = -\frac{1}{(\epsilon_0)^{1/2}} \sum_q \dot{e}_q(t) \mathbf{u}_q(\mathbf{R}) \quad (2.3.37a)$$

$$\mathbf{H}(\mathbf{R}, t) = \frac{1}{\mu_0(\epsilon_0)^{1/2}} \sum_q e_q(t) \nabla \times \mathbf{u}_q(\mathbf{R}) \quad (2.3.37b)$$

In summary, the electromagnetic field is dynamically equivalent to an infinite set of independent harmonic oscillators. A discussion of the physical implications of this result is postponed until Appendix L.

With this, I end the introductory discussion of the classic theory of electromagnetic radiation.

## Further Reading

The number of texts and treatises on Maxwell's equations is extremely large. A few of the many important texts are listed below. The reader is referred to these texts for further information on this subject.

1. J. D. Jackson, *Classical Electrodynamics*, 2nd ed., Wiley, New York, 1975.
2. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed., Pergamon Press, New York, 1975.
3. W. K. H. Panofsky and M. Phillips, *Classical Electricity and Magnetism*, 2nd ed., Addison-Wesley, Reading, MA, 1962.
4. P. Lorrain and D. R. Corson, *Electromagnetic Fields and Waves*, 2nd ed., Freeman, San Francisco, 1970.
5. H. C. Chen, *Theory of Electromagnetic Waves*, McGraw-Hill, New York, 1983.
6. M. Born and E. Wolf, *Principles of Optics*, 6th ed., Pergamon Press, New York, 1980.
7. F. A. Jenkins and H. E. White, *Fundamentals of Optics*, McGraw-Hill, New York, 1976.
8. R. M. A. Azzam, *Opt. Commun.* **41**, 1982, 223.
9. J. A. Ogilvy, *Theory of Wave Scattering from Random Rough Surfaces*, Adam Hilger, Bristol, 1991.
10. O. Wiener, *Ann. Phys.* **40**, 1890, 203.
11. J. Rosen, *Am. J. Phys.* **41**, 1973, 586.
12. R. H. Muller, *Surf. Sci.* **16**, 1969, 14.
13. W. H. Louisell, *Quantum Statistical Properties of Radiation*, Wiley, New York, 1973.

## PART 3

---

# POLARIZATION AND THE RADIATION FIELD

In Part 2, we discussed why the electric vector of electromagnetic waves is chosen as the variable of the theory, to characterize the polarization of the wavefield at a particular point in space.<sup>1</sup> In this part we shall examine the theoretical description of the electric vector and how it can cope with the analysis of partial polarization. The fields that one frequently deals with in optics exhibit some randomness. The very chaotic nature of the light emission process requires a statistical description of high-frequency electromagnetic radiation. One cannot study the rapid time fluctuations of the field but instead measure the correlations of the field at different space-time points. In its broadest sense, polarization is concerned mainly with the temporal correlation between two orthogonal components of the electric field vector to the direction of the wave propagation at a given point in space.

The plan of this part is as follows. In Section 3.1 we summarize a number of relevant definitions and basic concepts underlying the theory of partial polarization. Over the years, a number of theoretical ideas have shaped our understanding of this subject. But since the 1960s, it is firmly established that the coherency matrix provides the basis for the analysis of the correlations in optical fields. Anticipating further results, it is worth noting that the theory of partial polarization is intimately connected with the theory of partially coherent optical fields. In this chapter, we introduce several representations of the electric field vector and descriptive parameters of polarization in a form convenient for material to follow in later chapters. I have tried to include enough relevant detail so that the reader will not view the matter as a series of dogmatic statement. In Section 3.2, we encounter several geometric aspects of polarization through the use of a Poincaré sphere, which is an efficient graphical aid to represent a partially polarized wave by a single point in a parametrized space. The one-to-one correspondence among all possible states of polarization as distinct points of the Poincaré sphere with associated

<sup>1</sup>Polarization phenomena are customarily described by the electric field vector  $\mathbf{E}(\mathbf{R},t)$ . However, for many purposes, including the study of light propagating in condensed matter systems, the polarization of the wave is typically characterized by the electric displacement vector  $\mathbf{D}(\mathbf{R},t)$ , which represents the interaction of the electric field with the material. For an electromagnetic field in a linear medium whose macroscopic properties do not depend on time, both formulations of partial polarization theory are similar from the mathematical point of view, and in many applications, in fact, one does not need to be specific about what the vector really does or does not represent.

complex numbers is also developed. The convex set of polarization states of a plane wave is described in detail. We also examine a derivation of the "half the solid angle" formula. A discussion of an analogy between polarization optics and spin dynamics in a magnetic field is further developed. Specifically, we exploit the two-valuedness that determines the convex set of polarization states that has the same geometric structure as the spin- $\frac{1}{2}$  system. Section 3.3 is devoted to a discussion of mathematical details about the statistical properties of light under the assumption that the optical field can be described by a Gaussian random process. Emphasis is placed on a second-order statistical description of a stationary plane wavefield in the quasimonochromatic approximation. Our primary concern here is the issue of determining the statistics of the amplitude and phases of the electric field. Our secondary concern is determining the statistics of the observable quantities of the field at optical frequencies, specifically, the Stokes parameters. The practically important question of characterizing unpolarized light is further addressed. Finally, in Section 3.4, we derive, in some detail, formulas of the radiation entropy and reinterpret some of the results in earlier sections in conjunction with a useful analogy between the description of partially polarized wave and Ising spin systems. As a general comment, we emphasize that the description of the optical field proposed in this chapter is presented in completely classic terms. However, the discussion can also be given in a fully quantum-mechanical framework, as well. The connection between the quantization of harmonic oscillators and that of fields, which is a prerequisite in quantum electrodynamics, is briefly covered in Appendix L. The language of optics will be used, but the principles and equations presented are quite general and hold for all electromagnetic frequencies as long as the basic assumptions of the theory are met.

## SECTION 3.1

---

# Elementary Concepts and Definitions

This first chapter is intended to act as an introduction to the mathematical formalism required in later chapters. A secondary purpose is to establish notation and terminology. It is organized as follows. Before we describe the mathematical treatment, it is useful to examine the problem heuristically, particularly in order to point out that three different timescales characterize the temporal fluctuations that play an essential role in the theory of partial polarization. Next we concentrate on the description of the electric vector in terms of the complex analytic signal that plays a fundamental role in most analytic studies of partial polarization. The third section deals more specifically on the second-order statistical description of the field. Of particular importance, insofar as partial polarization is concerned, are the coherency matrix formalism and the Stokes parametrization of the electromagnetic field. We shall assume at the outset that the reader has been exposed previously to the theory of random processes.

### 3.1.1. TEMPORAL FLUCTUATIONS IN LIGHT BEAMS: CHARACTERISTIC TIMESCALES AND PARTIAL POLARIZATION

Light emitted from thermal radiations and laser sources fluctuates both spatially and temporally. For the usual thermal light sources, spatial fluctuations originate from the finite size of the source and to the spatial incoherence of the waves emitted from the different parts of the source. Temporal fluctuations are due to the wide blackbody spectrum emitted by such a source.

In Part 2, we restricted the discussion to monochromatic plane electromagnetic waves. A strictly monochromatic light, in the sense of infinitely long wavetrains, is completely polarized. In any practical experiment, the optical wave encountered can be described as a monochromatic radiation, but in fact this turns out to be a remote abstraction from reality because it requires the wave to be of infinite extent, that is, of zero spectral width. A real physical source has a finite-frequency bandwidth, although the frequency spectrum may be rather narrow. Here we shall be concerned mostly with quasimonochromatic waves; we assume that the bandwidth of the light, that is, the width of its power spectrum  $\Delta\nu$ , is very small compared with the central frequency  $\nu_0$ .

$(\Delta\nu \ll v_0)$ . A quasimonochromatic signal may result from the superposition of a large number of randomly timed statistically independent pulses with the same central frequency. The concept of partial polarization becomes significant for such quasimonochromatic wave.

There are three different characteristic timescales to consider in the analysis of partial polarization of quasimonochromatic light. The three timescales characterizing the fluctuations of the amplitude and phase of the electric vibrations are:  $\tau_1 \equiv 1/v_0$ , the period of the quasimonochromatic light, which is the time duration for the wave to go through one complete cycle;  $\tau_2 \equiv 1/\Delta\nu$ , the time of coherence, which provides a measure of the time interval over which the field behaves approximately monochromatic; and  $\tau_3$ , the integration response time of the detector (e.g., photomultiplier, photodiode).

Over time intervals long compared with  $\tau_2$ , the amplitude and phase of the electric field vector must be considered as random functions of time, while for intervals short compared with  $\tau_2$ , these quantities remain essentially constant. During times that are much smaller than  $\tau_2$ , natural light may be considered as completely polarized and described by an ellipse whose size, eccentricity, and orientation are randomly changed in time and whose orthogonal components are mutually correlated. In practice, actual measurements are exceedingly slow with respect to  $\tau_1$ ;  $\tau_3$  is so long that it integrates over many cycles of optical frequencies. In such cases a time average over many cycles of the field is performed.

Typical values of these timescales for narrowband lightwaves are on the order of  $\tau_1 (10^{-15} \text{ s}) \ll \tau_2 (10^{-9} - 10^{-4} \text{ s}) \ll \tau_3 (10^{-3} \text{ s})$ . For a thermal light source with a high degree of monochromaticity, such as a mercury arc lamp,  $\tau_2$  is of the order of  $10^{-8} \text{ s}$ . A typical HeNe laser with wavelength 632.8 nm has a Doppler-broadened spectral width of about 1.5 GHz corresponding to  $\tau_2 \cong 10^{-9} \text{ s}$ . For a well-stabilized laser, the coherence time of light can be as small as  $10^{-4} \text{ s}$ .

Many techniques have been widely exploited to measure  $\tau_2$ ; an example is Mach-Zehnder interferometry. For a detailed discussion of the different kind of measurements, see Refs 1 and 2.

### 3.1.2. THE ANALYTIC SIGNAL REPRESENTATION

Let us consider an optical field in the form of plane waves propagating in vacuum in the direction  $x_3$  characterized by the unit vector  $\mathbf{e}_3$ . For a light that propagates in a fixed direction, the electric vector is defined in a two-dimensional space transverse to that direction. Let  $E_j^{(r)}(\mathbf{R}, t)$  be the real-valued components of the electric vector at a point specified by position vector  $\mathbf{R}$  in the Euclidean space with basis  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ , at time  $t$ . The *state of polarization* refers to the electric field behavior at a particular point in space and, in general, varies from point to point of the field.

The real transverse field which obeys Maxwell's equations can be expressed as

$$\mathbf{E}^{(r)} = E_1^{(r)} \mathbf{e}_1 + E_2^{(r)} \mathbf{e}_2 \quad (3.1.1)$$

where  $\mathbf{e}_1, \mathbf{e}_2$  are two fixed orthogonal real unit vectors ( $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$ , where  $i, j = 1, 2$ ) in a plane normal to  $\mathbf{e}_3$  ( $\mathbf{e}_i \cdot \mathbf{e}_3 = 0$ , where  $i = 1, 2$ ). The subsequent analysis will deal with the complex analytic signal representation of the electric field rather than directly with the real electric vector [1-3]. This signal representation offers some unique and useful properties with a firm mathematical foundation as well as a number of physical advantages. On one hand, it involves only positive frequencies. On the other hand, one is led to it from a natural generalization to a nonmonochromatic field of the method of associating an exponential with a cosine, or a sine, as was done for a strictly monochromatic field in Part 2. Moreover, this approach has a useful counterpart in the quantum description of the electromagnetic field (Appendix L). The complex analytic signal associated with a real function is uniquely defined so that its real part is equal to the real function while its imaginary part is in quadrature with its real part. Because the analytic signal representation will prove to be very useful in the theory of partial polarization, we now devote a short discussion to some of its properties.

We shall assume that the field fluctuates in time and space, then  $E_1^{(r)}$  and  $E_2^{(r)}$  in Eq. (3.1.1) are random processes that may be regarded as typical members of an ensemble consisting of all possible realizations of the electric vector, which we shall assume to be stationary (at least in the wide sense). Moreover the stochastic field is as a rule treated as ergodic<sup>2</sup>. It is worth remarking that these conditions are likely to be encountered in practical problems. Since the field is a stationary random process, it is not square-integrable and hence its sample functions do not possess Fourier transforms. Such impossibility arises from the invariance under the translation of the origin of time of the probability distributions that characterize the statistical behavior of such a process; consequently its realizations cannot tend to zero as the time  $t \rightarrow \pm \infty$ . A technique for rendering a stationary random function such as  $\mathbf{E}(\mathbf{R}, t)$  Fourier transformable is to resort to truncated functions

$$E_j^{(r)}(\mathbf{R}, t; T) = \begin{cases} E_j^{(r)}(\mathbf{R}, t) & \text{when } |t| \leq T \\ 0 & \text{otherwise} \end{cases} \quad (3.1.2)$$

where  $T$  is some long time interval (i.e.,  $T \gg \tau_2$ ), that we may assume  $T \rightarrow \infty$ .

<sup>2</sup>For an ergodic random process any average over the entire ensemble of realizations may be obtained by a time averaging for just a single realization. Ensemble averages are performed on different realizations by repeating the experiment many times by using the same procedure for preparing the field. The problem of relating ensemble averages to time averages of a single realization is known as the *ergodic problem*. For discussion of the ergodic problem, the reader can consult Refs 4-6.

This condition is well satisfied in all problems of interest to us in the future [7]. The truncated functions are useful because they are square-integrable and thence are amenable to Fourier analysis.

$$E_j^{(r)}(\mathbf{R},t;T) = \int_{-\infty}^{\infty} \hat{E}_j^{(r)}(\mathbf{R},v;T) \exp(-2\pi i vt) dv \quad (3.1.3)$$

Our convention is that the caret means the Fourier transform on the time variable. Since  $E_j^{(r)}$  is real, the negative frequencies do not contain more physical information than the positive frequencies [i.e.,  $(\hat{E}_j^{(r)}(\mathbf{R},v;T))^* = \hat{E}_j^{(r)}(\mathbf{R},-v;T)$ ], we can rewrite Eq. (3.1.3) as

$$E_j^{(r)}(\mathbf{R},t;T) = 2\operatorname{Re} \left( \int_0^{\infty} \hat{E}_j^{(r)}(\mathbf{R},v;T) \exp(-2\pi i vt) dv \right) \quad (3.1.4)$$

In writing Eq. (3.1.4) we double the strength of the positive frequency components and remove the negative frequency components.

Now if we set

$$\hat{E}_j(\mathbf{R},v;T) \equiv a_j(\mathbf{R},v;T) \exp(i\theta_j(\mathbf{R},v;T)) \quad (3.1.5)$$

where  $a_j(\mathbf{R},v;T)$  and  $\theta_j(\mathbf{R},v;T)$  modulo  $(2\pi)$  are real functions, we arrive at the relation

$$E_j^{(r)}(\mathbf{R},t;T) = 2 \int_0^{\infty} a_j(\mathbf{R},v;T) \cos(-2\pi vt + \theta_j(\mathbf{R},v;T)) dv \quad (3.1.6)$$

which is the Fourier cosine integral representation of the real-valued signal  $E_j^{(r)}(\mathbf{R},v;T)$ . Similarly, one can introduce a Fourier sine integral function

$$E_j^{(i)}(\mathbf{R},t;T) = 2 \int_0^{\infty} a_j(\mathbf{R},v;T) \sin(-2\pi vt + \theta_j(\mathbf{R},v;T)) dv \quad (3.1.7)$$

by changing the phase of each spectral component of  $E_j^{(r)}$  by  $\pi/2$ . Making use of Eqs. (3.1.6) and (3.1.7), we can define the complex analytic signal  $E_j(\mathbf{R},t;T)$  associated with the real function  $E_j^{(r)}(\mathbf{R},t;T)$  as

$$E_j(\mathbf{R},t;T) \equiv E_j^{(r)}(\mathbf{R},t;T) + iE_j^{(i)}(\mathbf{R},t;T) \quad (3.1.8)$$

It follows that  $E_j(\mathbf{R},t;T)$  can be expressed in the form

$$E_j(\mathbf{R},t;T) \int_0^{\infty} \hat{E}_j(\mathbf{R},v;T) \exp(-2\pi i vt) dv \quad (3.1.9)$$

with

$$E_j(\mathbf{R}, v; T) \equiv 2 E_j^{(r)}(\mathbf{R}, v; T).$$

By Fourier inversion of Eq. (3.1.9), we have

$$\hat{E}_j(\mathbf{R}, v; T) = \begin{cases} \int_{-\infty}^{\infty} E_j(\mathbf{R}, v; T) \exp(2i\pi vt) dt & \text{when } v > 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.1.10)$$

The reason for the term *analytic* arises from the fact that the absence of negative frequency components in Eq. (3.1.9) ensures that each component of the realization  $\mathbf{E}(\mathbf{R}, t; T)$  considered as a function of  $t$ , will be analytic (i.e., free of singularities) and regular in the upper half of the complex plane.

So far no assumption has been made about the specific form of the power spectral density, which may be defined by the usual formula

$$g(v) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T |\hat{E}_j(\mathbf{R}, v; T)|^2 dt,$$

except that being concerned with quasimonochromatic waves  $g(v)$  is zero outside a narrow frequency range of width  $\Delta v$  that is small compared to the center frequency  $v_0$ . A basic property of the spectral density follows from Parseval's theorem:

$$\begin{aligned} & \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (E_j^{(r)}(\mathbf{R}, t; T))^2 dt \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T (E_j^{(i)}(\mathbf{R}, t; T))^2 dt = 2 \int_0^{\infty} g(v) dv \end{aligned} \quad (3.1.11)$$

Two spectral profiles are important to us for later purposes:

$$\text{Gaussian spectrum: } g(v) \sim \frac{1}{\Delta v} \exp \left[ -\left( \frac{v - v_0}{\Delta v} \right)^2 \right] \quad (3.1.12a)$$

$$\text{Lorentzian spectrum: } g(v) \sim \frac{\Delta v}{[(v - v_0)^2 + (\Delta v)^2]} \quad (3.1.12b)$$

For completeness, we indicate two alternative parametrizations of complex analytic signals. First,  $E_j^{(r)}$  and  $E_j^{(i)}$  are conjugate functions and are related by

Hilbert's reciprocity relations

$$E_j^{(i)}(\mathbf{R},t;T) = \frac{1}{\pi} \text{PV} \left( \int_{-\infty}^{\infty} \frac{E_j^{(r)}(\mathbf{R},t';T)}{t' - t} dt' \right) = \mathbf{H}(E_j^{(r)}(\mathbf{R},t;T)) \quad (3.1.13a)$$

$$E_j^{(r)}(\mathbf{R},t;T) = -\frac{1}{\pi} \text{PV} \left( \int_{-\infty}^{\infty} \frac{E_j^{(i)}(\mathbf{R},t';T)}{t' - t} dt' \right) = -\mathbf{H}(E_j^{(i)}(\mathbf{R},t;T)) \quad (3.1.13b)$$

where  $\mathbf{H}$  is the Hilbert transform and the symbol PV means that the principal value in Cauchy's sense of the integral must be taken at  $t' = t$ .<sup>3</sup> These relations enable us in principle to calculate  $E_j^{(i)}(\mathbf{R},t;T)$  once we know  $E_j^{(r)}(\mathbf{R},t;T)$  and vice versa. This fact, as well as other extensions, are discussed extensively in Refs. 7 and 8.

The truncated function may be also rewritten as

$$E_j(\mathbf{R},t;T) \equiv E_j^{(r)}(\mathbf{R},t;T) + i\mathbf{H}(E_j^{(r)}(\mathbf{R},t;T)) \quad (3.1.14)$$

and the complex  $\mathbf{E}(\mathbf{R},t)$  is obtained from  $E_j(\mathbf{R},t;T)$  by letting  $T \rightarrow \infty$ . A second expression of the analytic signal can be obtained in terms of the improper function  $\delta_-(t' - t)$  and using the well-known identity

$$\delta_-(t' - t) = \int_0^\infty \exp(2i\pi v(t' - t)) dv = \frac{1}{2} \left( \delta(t' - t) + \frac{i}{\pi} \text{PV} \left( \frac{1}{t' - t} \right) \right) \quad (3.1.15)$$

where  $\delta(t)$  denotes the Dirac delta function. Using Eq. (3.1.15), we find an expression of the analytic signal of the electric field that reads as a linear transform of its real-valued component

$$E_j(\mathbf{R},t;T) = 2 \int_{-\infty}^{\infty} E_j^{(r)}(\mathbf{R},t';T) \delta_-(t' - t) dt' \quad (3.1.16)$$

This equation shows that  $E_j(\mathbf{R},t;T)$  is a linear transform of  $E_j^{(r)}(\mathbf{R},t';T)$ . It is assumed here that not only the real fields  $\mathbf{E}^{(r)}$  but also the complex analytic signal  $\mathbf{E}$  satisfy Maxwell's equations. This can be shown by Hilbert transforming Maxwell's equations for the real field  $\mathbf{E}^{(r)}$  and using the fact that if two functions are Hilbert transforms of each other, so are their derivatives. One then finds that  $\mathbf{E}^{(i)}$  and hence the complex analytic signal  $\mathbf{E} = \mathbf{E}^{(r)} + i\mathbf{E}^{(i)}$  obey the Maxwell equations and furthermore satisfy whatever boundary conditions we require of the electric field vector, such as periodic boundary conditions.

Note that the complex analytic signal associated with the truncated magnetic vector can be written in a similar manner to the electric vector as  $H_j(\mathbf{R},t;T) \equiv H_j^{(r)}(\mathbf{R},t;T) + iH_j^{(i)}(\mathbf{R},t;T)$ , where  $H_j^{(r)}(\mathbf{R},t;T)$  and  $H_j^{(i)}(\mathbf{R},t;T)$  are respectively the real and imaginary parts of the complex magnetic field vector. In subsequent formulas, we drop explicit reference to the variable  $T$ , except when this would lead to ambiguity.

<sup>3</sup>Under the assumption that the wave forms a narrow band about a large central frequency  $v_0$ , we note that  $E_j^{(i)}(\mathbf{R},t) \cong E_j^{(r)}(\mathbf{R},t + (1/4v))$ . The expression for  $E_j^{(i)}(\mathbf{R},t)$  is only an approximation that depends on the quasimonochromatic character of the wave.

A note about terminology: One customarily writes

$$E_j^{(r)}(\mathbf{R}, t) \equiv a_j(\mathbf{R}, t) \cos(\theta_j(\mathbf{R}, t) - 2\pi v_0 t) \quad (3.1.17a)$$

$$E_j^{(i)}(\mathbf{R}, t) \cong a_j(\mathbf{R}, t) \sin(\theta_j(\mathbf{R}, t) - 2\pi v_0 t) \quad (3.1.17b)$$

Note that the following approximation for the imaginary part of the complex analytic signal can be made with good accuracy (refer back to footnote 3).

In fact, we do not need the full expression of the wave given by Eqs. (3.1.17a-b). By suppressing the harmonic dependence, we obtain the two-dimensional complex column vector  $\mathbf{E}(\mathbf{R})$ , which specifies the polarization state of the wave at a fixed point  $\mathbf{R}$  in space:

$$\mathbf{E}(\mathbf{R}) = \begin{bmatrix} a_1(\mathbf{R}) \exp(i\theta_1(\mathbf{R})) \\ a_2(\mathbf{R}) \exp(i\theta_2(\mathbf{R})) \end{bmatrix} = \exp(i\theta_1(\mathbf{R})) \begin{bmatrix} a_1(\mathbf{R}) \\ a_2(\mathbf{R}) \exp(i\theta(\mathbf{R})) \end{bmatrix} \quad (3.1.18)$$

In subsequent formulas we shall drop the spatial information about the wave by considering the field over the fixed transverse plane  $x_3 = 0$ . This spatial term can be dropped because the electric field vector has the same phase at all points on a  $x_3 = \text{constant}$  plane. Moreover, we assume that the field is uniform so will be independent of  $x_1$  and  $x_2$ . Then, apart from an unessential constant multiplicative phase factor, Eq. (3.1.18) permits us to write

$$\mathbf{E} = \begin{bmatrix} a_1 \\ a_2 \exp(i\theta) \end{bmatrix} \quad (3.1.19)$$

This vector, called the *Jones vector of the plane uniform wave*, contains information about the amplitudes and phases of the field. To facilitate the analysis, it is often more convenient to consider the normalized Jones vector, which is defined by  $\mathbf{E}/\sqrt{a_1^2 + a_2^2}$ .

Thus far we have considered a real rectangular basis. In some problems, it is more appropriate to use the complex circularly polarized representation of the transverse field (consisting of the right-handed vector  $\mathbf{e}_r \equiv 1/\sqrt{2}(\mathbf{e}_1 - i\mathbf{e}_2)$  and of the left-handed vector  $\mathbf{e}_l \equiv 1/\sqrt{2}(\mathbf{e}_1 + i\mathbf{e}_2)$ ). The complex vectors  $\mathbf{e}_r$  and  $\mathbf{e}_l$  satisfy the orthonormality relations  $\mathbf{e}_r^* \cdot \mathbf{e}_r = \mathbf{e}_l^* \cdot \mathbf{e}_l = 1$  and  $\mathbf{e}_r^* \cdot \mathbf{e}_l = \mathbf{e}_l^* \cdot \mathbf{e}_r = 0$ . The complex basis vectors  $(\mathbf{e}_l, \mathbf{e}_r)$  form the helicity basis. Then, under the change of basis, the Jones vectors will undergo the unitary transformations<sup>4</sup>:

$$\mathbf{E}_c = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -i \end{bmatrix} \mathbf{E} = \mathbf{U}_c \mathbf{E} \quad (3.1.20a)$$

$$\mathbf{E} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -i & i \end{bmatrix} \mathbf{E}_c = \mathbf{U}_c^{-1} \mathbf{E}_c \quad (3.1.20b)$$

where the subscript  $c$  indicates helicity basis.

<sup>4</sup>The general unitary matrix satisfies the requirement  $\mathbf{U}^* \equiv \mathbf{U}^{T*} = \mathbf{U}^{-1}$ , where the superscript  $T$  denotes the transpose.

### 3.1.3. THE INSTANTANEOUS ELLIPSE FOR THE ELECTRIC FIELD

This section deals with the mathematical description of the ellipse for the electric field of a plane electromagnetic uniform light beam propagating in the positive direction of the  $x_3$  axis of a right-handed Cartesian coordinate system ( $x_1, x_2, x_3$ ), at a fixed point  $\mathbf{R}$  in space. By a suitable choice of reference frame, we consider that this point of observation is localized within the plane  $x_3 = 0$ . We also assume that the field possesses a narrowband power spectrum of width  $\Delta\nu \ll v_0$  which is its center frequency.

Let  $E_j^{(r)}(t)$  (where  $j = 1, 2$ ) denote the real orthogonal components of the complex electric field vector

$$E_1^{(r)}(t) \equiv a_1(t) \cos(\theta_1(t) - 2\pi v_0 t) \quad (3.1.21a)$$

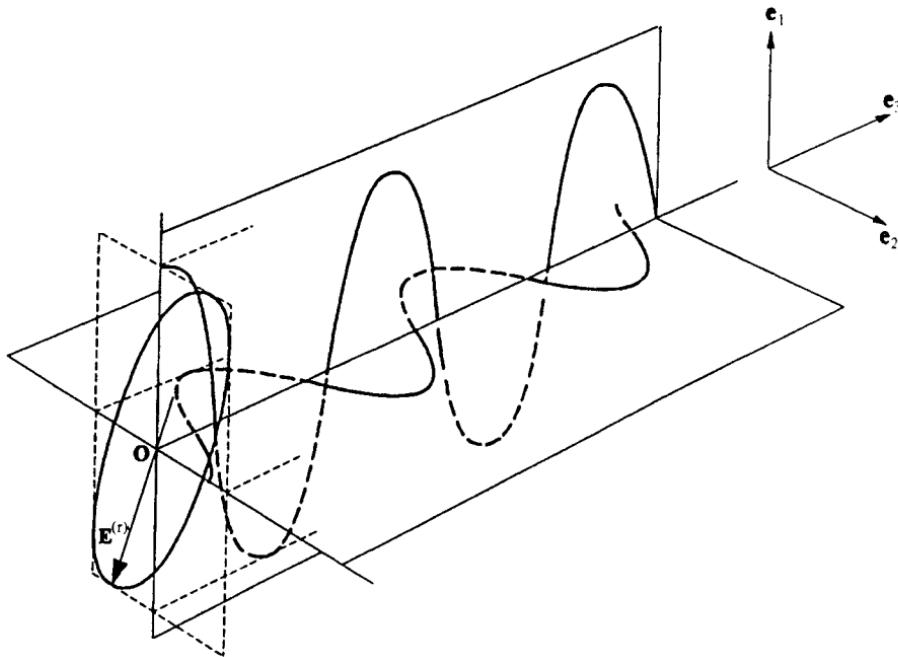
$$E_2^{(r)}(t) \equiv a_2(t) \cos(\theta_2(t) - 2\pi v_0 t) \quad (3.1.21b)$$

where  $a_1(t)$  and  $a_2(t)$  are the instantaneous amplitudes (envelopes) along the  $x_1$  and  $x_2$  axes, and  $\theta_1(t)$  and  $\theta_2(t)$  represent the respective instantaneous phases at a fixed point in space as a function of time, respectively. These signals fluctuate slowly in comparison with  $\cos(2\pi v_0 t)$  at optical frequencies, that is, they are nearly constant over any interval short compared to  $\tau_2$ . As the quasimonochromatic wave propagates through space, we find that, in a fixed plane ( $\mathbf{e}_1, \mathbf{e}_2$ ) perpendicular to the light's direction of travel ( $\mathbf{e}_3$ ), the end point of the electric vector at a fixed point in space traces out an ellipse, whose shape changes continuously. This is illustrated in Fig. 3.1.1.

When the ellipse maintains a constant orientation, ellipticity, and sense in which the ellipse is described, the wave is said to be completely polarized at that point. Therefore, completely polarized light is, in general, elliptically polarized apart several degenerate forms of the ellipse. At the other extreme, if the successive ellipses traced out over many periods of oscillations of the electric field exhibit no regular pattern, the wave is said to be unpolarized. Between these two extreme cases, the wave is said to be partially polarized. Such an ellipse is periodically described in time at a rate  $(\tau_1)^{-1}$ . An instantaneous snapshot of the electric vector would show a flattened helix in space; the field vector rotates and, at the same time, changes in magnitude. Note that the amplitude of the optical field, and by way of consequence, the tracing of the ellipse is not an observable at optical frequencies.

A sample realization of the locus of the endpoint of the electric field vector traced out by  $E_1^{(r)}(t)$  and  $E_2^{(r)}(t)$  at a fixed point in space as a function of time can be obtained by eliminating  $2\pi v_0 t$  between  $E_1^{(r)}(t)$  and  $E_2^{(r)}(t)$  into Eqs. (3.1.21a–b). We therefore write the equation of the polarization ellipse of the electric field.

$$\left(\frac{E_1^{(r)}(t)}{a_1(t)}\right)^2 + \left(\frac{E_2^{(r)}(t)}{a_2(t)}\right)^2 - 2 \left(\frac{E_1^{(r)}(t)E_2^{(r)}(t)}{a_1(t)a_2(t)}\right) \cos(\theta(t)) = (\sin(\theta(t)))^2 \quad (3.1.22)$$



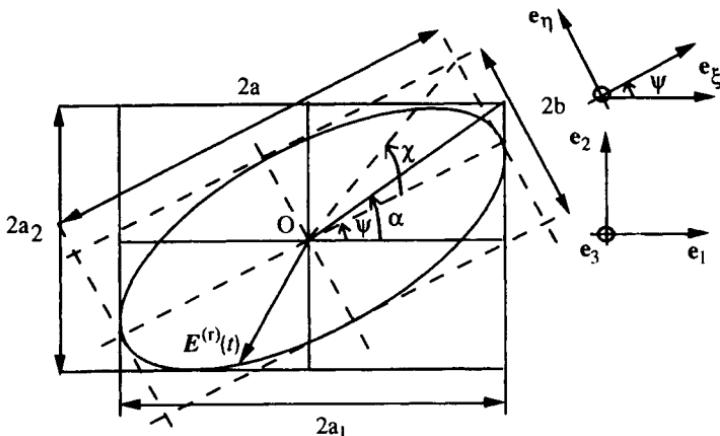
**Figure 3.1.1.** The polarization ellipse of the electric field vector. At any point along a beam of polarized light, the endpoint of the electric field vector traces an ellipse in a plane perpendicular to the beam.

where the phase angle  $\theta(t) \equiv \theta_2(t) - \theta_1(t)$  is defined in the range  $-\pi \leq \theta < \pi$ . It is important to note that, in fact,  $\theta_1(t)$ ,  $\theta_2(t)$ , and  $\theta(t)$  are angles reduced modulo  $2\pi$  since they appear only in the arguments of trigonometric functions. When  $\theta = \pi/2$ , the reference axes coincide with the major and minor axes of the ellipse. In all other cases the reference axes do not correspond to the major and minor axes of the ellipse. Let us consider Fig. 3.1.2 in order to illustrate the elementary approach we are discussing. The ellipse defined by Eq. (3.1.22) is inscribed in a rectangle whose lengths are  $2a_1$  and  $2a_2$  (Fig. 3.1.2). The electric field vector is said to be elliptically polarized.

For monochromatic light, the amplitudes and phases are constant, so Eq. (3.1.22) reduces to:

$$\left(\frac{E_1^{(r)}(t)}{a_1}\right)^2 + \left(\frac{E_2^{(r)}(t)}{a_2}\right)^2 - 2 \left(\frac{E_1^{(r)}(t)E_2^{(r)}(t)}{a_1a_2}\right) \cos(\theta) = (\sin(\theta))^2 \quad (3.1.23)$$

For a quasimonochromatic radiation, these quantities are slowly varying functions of time. Note that the cross-term  $E_1^{(r)}(t)E_2^{(r)}(t)$  in Eq. (3.1.23) implies that the polarization ellipse of the electric field is in general rotated through an angle  $\psi$  (see Fig. 3.1.2).



**Figure 3.1.2.** The different parameters that describe the geometry of the instantaneous polarization ellipse in the plane  $(\mathbf{e}_1, \mathbf{e}_2)$  at a given point in space for a transverse optical wave propagating along the direction characterized by unit vector  $\mathbf{e}_3$ :  $a$  is the length of the semimajor axis of the ellipse,  $b$  is the length of its semiminor axis,  $\chi$  is the ellipticity angle such that  $\tan(\chi) = \pm b/a$  where  $b/a$  is the ratio of the length of the semiminor to the length of the semimajor axis, and  $\psi$  is the azimuth angle that the major axis makes with the positive 1-axis. The angle  $\alpha$  is related to  $a_1$  and  $a_2$  by  $\tan(\alpha) = a_2/a_1$ .

As embodied by Eq. (3.1.23), four independent parameters are needed to completely characterize the ellipse: the amplitudes  $a_1$ ,  $a_2$  and the phase difference  $\theta$  (absolute value and sign). However, it is more convenient to consider the following set of parameters: (1) the intensity  $S_0$  ( $0 < S_0 < \infty$ ) governing the size of the ellipse; (2) the ellipticity angle  $\chi$ , which is defined by the arctangent of the ratio of the length of the semiminor axis to the length of the semimajor axis (3); the azimuth angle  $\psi$  ( $0 \leq \psi < \pi$ ), which the major axis makes with the positive 1-axis; and (4) the sense of rotation of the endpoint of the electric field (also termed the “handedness”). The notation  $S_0$  for the intensity of the wave will be developed later, in Section 3.3. We should point out that the two former parameters are independent of the system of coordinates chosen while the two latter parameters are to be referred to the coordination frame. We adopt here a conversion based on the sense of rotation of the endpoint of the electric vector as viewed by an observer receiving the radiation. The polarization is said to be right-handed [ $0 < \theta < \pi$ , i.e., if  $\sin(\theta) > 0$ ] when the electric vector rotates in a clockwise rotation and left-handed [ $-\pi < \theta < 0$ , i.e., if  $\sin(\theta) < 0$ ] in the opposite situation. It is important to note that this convention is linked to the choice of the time-dependent part of the exponential,  $\exp(-2\pi i vt)$ . If this was chosen to be  $\exp(2\pi i vt)$ , the handedness of the polarization would be reversed. The length of the semiminor axis  $a$  and the length of the semimajor axis  $b$  ( $a \geq b$ ) have

the property that

$$S_0 \equiv a_1^2 + a_2^2 = a^2 + b^2, \quad (3.1.24)$$

The problem is now to determine the shape and the orientation of the ellipse. Let  $x_\xi$  and  $x_\eta$  be a new set of axes (sometimes called the *principal axes*) along the major and minor axes of the ellipse, where  $\psi$  denotes the angle between  $x_\xi$  and the 1-axis (Fig. 3.1.2). The two sets of axes differ from each other by a spatial rotation:  $\theta_\eta - \theta_\xi = \pi/2$ . With respect to the principal coordinate system, the electric field components become

$$E_\xi^{(r)}(t) \equiv a \cos(\theta_\xi - 2\pi v t) \quad (3.1.25a)$$

$$E_\eta^{(r)}(t) \equiv b \cos(\theta_\eta - 2\pi v t) = \pm b \sin(\theta_\xi - 2\pi v t) \quad (3.1.25b)$$

so that we obtain the equation of an ellipse in standard form

$$\left(\frac{E_\xi^{(r)}(t)}{a}\right)^2 + \left(\frac{E_\eta^{(r)}(t)}{b}\right)^2 = 1 \quad (3.1.26)$$

Note that the  $\pm$  sign in Eq. (3.1.25b) differentiates between the two possible senses in which the endpoint of the electric field may rotate to describe the ellipse. The electric field components in the principal coordinate system  $E_\xi$  and  $E_\eta$  are obtained from those in the  $x_1$  and  $x_2$  coordinate system by the transformation

$$E_\xi = E_1 \cos(\psi) + E_2 \sin(\psi) \quad (3.1.27a)$$

$$E_\eta = -E_1 \sin(\psi) + E_2 \cos(\psi) \quad (3.1.27b)$$

We now place Eqs. (3.1.21a, b) and (3.1.25a, b) into Eqs. (3.1.27a, b). Thus

$$\begin{aligned} & a(\cos(\theta_\xi) \cos(2\pi v t) - \sin(\theta_\xi) \sin(2\pi v t)) \\ &= a_1(t) \cos(\theta_1(t) - 2\pi v t) \cos(\psi) + a_2(t) \cos(\theta_2(t) - 2\pi v t) \sin(\psi) \end{aligned} \quad (3.1.28a)$$

$$\begin{aligned} & \pm b(\sin(\theta_\xi) \cos(2\pi v t) + \cos(\theta_\xi) \sin(2\pi v t)) \\ &= -a_1(t) \cos(\theta_1(t) - 2\pi v t) \sin(\psi) + a_2(t) \cos(\theta_2(t) - 2\pi v t) \cos(\psi) \end{aligned} \quad (3.1.28b)$$

We can expand the cosine and sine terms in the right-hand side of Eqs. (3.1.28a, b) and equate the coefficients of  $\cos(2\pi v t)$  and  $\sin(2\pi v t)$ . The result is

$$a \cos(\theta_\xi) = a_1(t) \cos(\theta_1(t)) \cos(\psi) + a_2(t) \cos(\theta_2(t)) \sin(\psi) \quad (3.1.29)$$

$$a \sin(\theta_\xi) = a_1(t) \sin(\theta_1(t)) \cos(\psi) + a_2(t) \sin(\theta_2(t)) \sin(\psi) \quad (3.1.30)$$

$$\pm b(\sin(\theta_\xi)) = -a_1(t) \cos(\theta_1(t)) \sin(\psi) + a_2(t) \cos(\theta_2(t)) \cos(\psi) \quad (3.1.31)$$

$$\pm b(\cos(\theta_\xi)) = a_1(t) \sin(\theta_1(t)) \sin(\psi) - a_2(t) \sin(\theta_2(t)) \cos(\psi) \quad (3.1.32)$$

Squaring and adding Eqs. (3.1.29) and (3.1.30) and similarly from Eqs. (3.1.31) and (3.1.32), we obtain, after rearrangement of terms

$$a^2 = a_1^2 \cos^2(\psi) + a_2^2 \sin^2(\psi) + 2a_1 a_2 \cos(\theta) \cos(\psi) \sin(\psi) \quad (3.1.33a)$$

$$b^2 = a_1^2 \sin^2(\psi) + a_2^2 \cos^2(\psi) - 2a_1 a_2 \cos(\theta) \cos(\psi) \sin(\psi) \quad (3.1.33b)$$

Adding Eqs. (3.1.33a) and (3.1.33b), we obtain the following equations

$$a_1^2 + a_2^2 = a^2 + b^2 \quad (3.1.34a)$$

$$a_1^2 - a_2^2 = (a^2 - b^2) \cos(2\psi) \quad (3.1.34b)$$

On multiplying Eq. (3.1.29) by (3.1.31) and (3.1.30) by (3.1.32) and adding, we have

$$\pm ab = a_1 a_2 \sin(\theta) \quad (3.1.35a)$$

$$(a^2 - b^2) \sin(2\psi) = a_1 a_2 \cos(\theta) \quad (3.1.35b)$$

Next we divide Eq. (3.1.31) by (3.1.30) and (3.1.32) by (3.1.29) to get

$$\frac{\pm b}{a} = \frac{-a_1(t) \cos(\theta_1(t)) \sin(\psi) + a_2(t) \cos(\theta_2(t)) \cos(\psi)}{a_1(t) \sin(\theta_1(t)) \cos(\psi) + a_2(t) \sin(\theta_2(t)) \sin(\psi)} \quad (3.1.36a)$$

$$\frac{\pm b}{a} = \frac{-a_1(t) \cos(\theta_1(t)) \sin(\psi) + a_2(t) \cos(\theta_2(t)) \cos(\psi)}{a_1(t) \cos(\theta_1(t)) \cos(\psi) + a_2(t) \cos(\theta_2(t)) \sin(\psi)} \quad (3.1.36b)$$

After equating these two equations, we get

$$(a_1^2 - a_2^2) \sin(2\psi) = 2a_1 a_2 \cos(\theta) \cos(2\psi) \quad (3.1.37)$$

hence we arrive at the relation

$$\tan(2\psi) = \frac{2a_1 a_2}{a_1^2 - a_2^2} \cos(\theta) \quad (3.1.38)$$

The angle  $\psi$  can be also rewritten

$$\tan(2\psi) = (\tan(2\alpha)) \cos(\theta) \quad (3.1.39)$$

where we have defined an angle  $\alpha$  ( $0 \leq \alpha \leq \pi/2$ )

$$\tan(\alpha) = \frac{a_2}{a_1} \quad (3.1.40)$$

Similarly, after trigonometric algebra, we obtain

$$\sin(2\chi) = \frac{2a_1 a_2}{a_1^2 + a_2^2} \sin(\theta) = \sin(2\alpha) \sin(\theta) \quad (3.1.41)$$

where the angle  $\chi$  ( $-\pi/4 \leq \chi \leq \pi/4$ ) such that

$$\tan(\chi) = \frac{\pm b}{a} \quad (3.1.42)$$

specifies the shape and the orientation of the ellipse, with the plus sign for left-hand polarization and the minus sign for right-hand polarization. For completeness, we indicate the following relations between the parameters of the polarization ellipse

$$\cos(2\alpha) = \cos(2\chi) \cos(2\psi) \quad (3.1.43)$$

$$\tan(\theta) = \tan(2\chi)(\sin(2\psi))^{-1} \quad (3.1.44)$$

Equation (3.1.43) is obtained by dividing Eq. (3.1.34b) by Eq. (3.1.34a) and making use of Eqs. (3.1.40) and (3.1.42). Equation (3.1.44) follows from simple manipulation of the preceding equations.

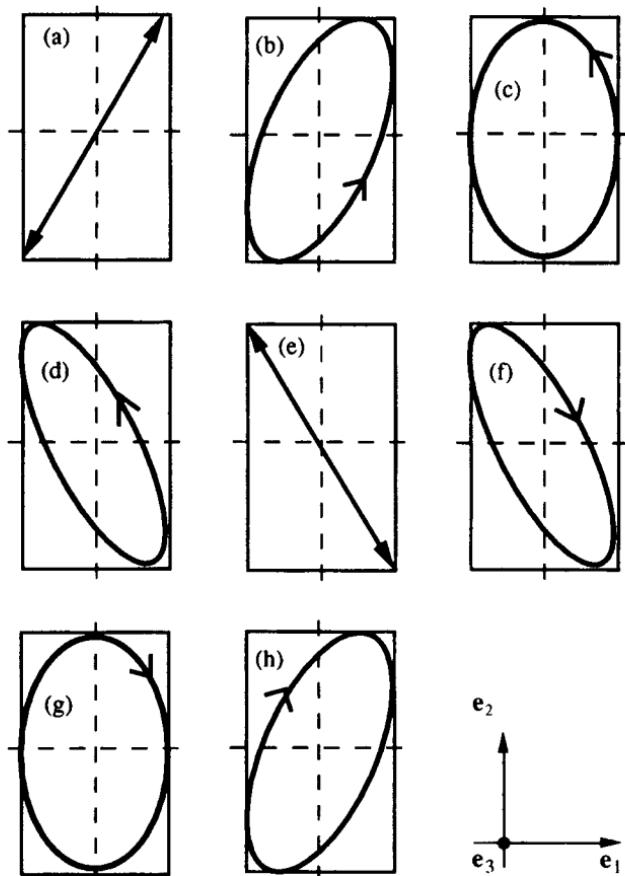
Figure 3.1.3 displays different Lissajous representations of polarization ellipses for different phase angles  $\theta$ . It may be useful to discuss several cases of interest. If  $\theta = p\pi$  ( $p = 0, \pm 1, \pm 2, \dots$ ),  $\mathbf{E}$  is linearly polarized. The ellipse reduces to a segment of straight line, and we have  $\chi = (-1)^p(a_2/a_1)$ . The Jones vectors are

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

for horizontal and vertical states, respectively. If  $\theta = p\pi/2$  ( $p = \pm 1, \pm 3, \dots$ ) and  $a_1 = a_2$ , then  $\mathbf{E}$  is circularly polarized. The ellipse degenerates into a circle, and we have  $\chi = \pm(\pi/4)$ . The state is right-handed if  $\sin(\theta) > 0$  corresponding to  $\chi = \pi/4$  and left-handed if  $\sin(\theta) < 0$  corresponding to  $\chi = -\pi/4$ . The Jones vectors of the helicity basis are

$$\mathbf{e}_l = 1/\sqrt{2} \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \text{and} \quad \mathbf{e}_r = 1/\sqrt{2} \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

The Jones vector of a general elliptic polarization state may be written (up to an absolute phase, which is usually not of interest, of modulus unity) in the



**Figure 3.1.3.** Lissajous diagrams of polarization ellipse showing the linear (a)  $\theta = 0$ , (e)  $\theta = \pi$ ; and elliptical (b)  $0 < \theta < \pi/2$ , (c)  $\theta = \pi/2$ , (d)  $\pi/2 < \theta < \pi$ , (f)  $\pi < \theta < 3\pi/2$ , (g)  $\theta = 3\pi/2$ , (h)  $3\pi/2 < \theta < 2\pi$  paths traced out by the tip of a rotating electric vector during one period of the wavefield. Note that the wave would be circularly polarized when  $\theta = \pi/2$  or  $3\pi/2$  if  $a_1 = a_2$ . The light is propagating out of the plane of the paper.

form

$$\mathbf{E}(\alpha, \theta) = \begin{bmatrix} \cos(\alpha) \\ \sin(\alpha) \exp(i\theta) \end{bmatrix} \quad (3.1.45)$$

In the principal coordinate system, the Jones vector is given by

$$\begin{bmatrix} \cos(\psi) \\ i \sin(\psi) \end{bmatrix}$$

Equations (3.1.27a, b) allow us to express the Jones vector in either a Cartesian

**TABLE 3.1.1. Special Degenerate Forms of Elliptically Polarized Waves Characterized by Normalized States of Polarization  $\mathbf{E}$  and Completely Unpolarized State, Their Corresponding Stokes Vector  $\mathbf{S}$ , Complex Polarization Ratio  $Z_{12}$ , and Density Matrix  $\mathbf{D}_2^a$**

Polarization	$\mathbf{E}$	$Z_{12}$	$\mathbf{E}_c$	$\mathbf{S}$	$\mathbf{D}_2$	Symbol
Linear horizontal	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	0	$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$	$\leftrightarrow$
Linear vertical	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\infty$	$\frac{1}{\sqrt{2}} \begin{bmatrix} i \\ -i \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$	$\uparrow$
Linear $+45^\circ$	$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$	1	$\frac{1}{2} \begin{bmatrix} 1 & +i \\ 1 & -i \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$	$\frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$	$\checkmark$
Linear $-45^\circ$	$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$	-1	$\frac{1}{2} \begin{bmatrix} 1 & -i \\ 1 & +i \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}$	$\frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$	$\checkmark$
Right-handed circular	$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix}$	$-i$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}$	$\frac{1}{2} \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$	$\odot$	
Left-handed circular	$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}$	$i$	$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}$	$\frac{1}{2} \begin{bmatrix} 1 & -i \\ i & -1 \end{bmatrix}$	$\odot$
Unpolarized				$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$\frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\ast$

<sup>a</sup>All quantities are expressed in the linear polarization basis. The term  $\mathbf{E}_c$  is the Jones vector expressed in the helicity basis. The pictorial representations of the polarization states to be used throughout this book are also listed. The time evolution of the tip of the electric field for unpolarized light is a random isotropic orientation distribution in a plane perpendicular to the light's direction of propagation.

coordinate system as

$$\mathbf{E}(\psi, \chi) = \begin{bmatrix} \cos(\psi) \cos(\chi) - i \sin(\psi) \sin(\chi) \\ \sin(\psi) \cos(\chi) + i \cos(\psi) \sin(\chi) \end{bmatrix} \quad (3.1.46a)$$

or a circular coordinate systems as

$$\mathbf{E}_c(\psi, \chi) = \frac{1}{\sqrt{2}} \begin{bmatrix} (\cos(\chi) - \sin(\chi)) \exp(i\psi) \\ (\cos(\chi) + \sin(\chi)) \exp(-i\psi) \end{bmatrix} \quad (3.1.46b)$$

Elliptically polarized light contains all linear polarizations in the sense that if it is passed through a linear polarizer, then some linearly polarized light will always be transmitted. This fact is developed in greater detail in Section 4.1.

The condition that two Jones vectors are orthogonal is

$$\mathbf{E}_a^* \cdot \mathbf{E}_b = \mathbf{E}_a \cdot \mathbf{E}_b^* = 0 \quad (3.1.47)$$

The pair of normalized Jones vectors  $\mathbf{E}_a$  and  $\mathbf{E}_b$  is called an *orthonormal pair of vectors*, for example,  $(\mathbf{e}_r, \mathbf{e}_l)$ . The ellipse of polarization corresponding to a Jones vector orthonormal to a Jones vector characterized by azimuth  $\psi$  and ellipticity  $\chi$  has equal and opposite ellipticity, (i.e.,  $\chi \rightarrow -\chi$ ), and its major axis is orthogonal (i.e.,  $\psi \rightarrow \psi \pm \pi/2$ ). The ellipses of polarization that correspond to two orthogonal states are traced in opposite senses. It is also noteworthy that linearly polarized states are such that

$$\mathbf{E}^* = \mathbf{E} \quad (3.1.48)$$

while circularly polarized states have the property

$$\mathbf{E} \cdot \mathbf{E} = 0 \quad (3.1.49)$$

A number of quantities have been defined in this section. As an aid to the reader, we summarize the relevant parameters in Table 3.1.1.

### 3.1.4. THE COMPLEX POLARIZATION RATIO REPRESENTATION OF POLARIZED LIGHT

The foregoing discussion was based on the description of the wave by the Jones vector formalism. However, for many purposes (e.g., ellipsometry), it is more appropriate, from the physical point of view, to characterize the wave by an alternative formulation, stated in terms of the complex polarization ratio. The complex variable  $Z_{12}$  is defined by the ratio of the two orthogonal components of the Jones vector related to the linear polarization basis  $(\mathbf{e}_1, \mathbf{e}_2)$ :

$$Z_{12} \equiv \frac{E_2}{E_1} = \frac{a_2}{a_1} \exp(i(\theta_2 - \theta_1)) = \tan(\alpha) \exp(i\theta) = \text{Re}(Z_{12}) + i \text{Im}(Z_{12}) \quad (3.1.50)$$

A number of properties of  $Z_{12}$  follow from Eq. (3.1.50). This representation provides a convenient way to specify the state of polarization of a completely polarized plane wave of infinite extent (see Table 3.1.1). We illustrate this parametrization in two cases. For a left-handed circular polarization,  $a_1 = a_2$  and  $\sin(\theta) = 1$ , and according to Eq. (3.1.50),  $Z_{12} = i$ . Similarly, we find a polarization ratio  $Z_{12} = -i$  for a right-handed circular polarization. The linear horizontal polarization state corresponds to the ratio  $Z_{12} = 0$ , and the orthogonal vertical polarization state corresponds to  $Z_{12} = \pm\infty$ . Each polarization state can be uniquely represented by its complex variable  $Z_{12}$ . In this basis, both the azimuth  $\psi$  and ellipticity  $\chi$  angles can be obtained from  $Z_{12}$  by the following relations [9]:

$$\tan(2\psi) = \frac{2\operatorname{Re}(Z_{12})}{1 - |Z_{12}|^2} \quad (3.1.51a)$$

$$\sin(2\chi) = \frac{2\operatorname{Im}(Z_{12})}{1 + |Z_{12}|^2} \quad (3.1.51b)$$

Conversely, the polarization ratio can also be written in terms of the azimuth  $\psi$  and ellipticity  $\chi$  angles as

$$Z_{12} = \frac{\tan(\psi) + i\tan(\chi)}{1 - i\tan(\psi)\tan(\chi)} \quad (3.1.52)$$

This is a most remarkable fact. It asserts that each elliptically polarized wave of azimuth  $\psi$  and ellipticity  $\chi$  can be parametrized by a single complex number. We defer a full discussion of the geometric interpretation of Eq. (3.1.52) and the implications of the one-to-one correspondence among all the polarization ratios and the corresponding points on the Poincaré sphere. Note that we may use Eq. (3.1.50) to rewrite the angles  $\alpha$  and  $\theta$  in the form

$$(\tan \alpha)^2 = \operatorname{Re}^2(Z_{12}) + \operatorname{Im}^2(Z_{12}) \quad (3.1.53a)$$

$$\tan(\theta) = \frac{\operatorname{Im}(Z_{12})}{\operatorname{Re}(Z_{12})} \quad (3.1.53b)$$

To each complex number  $Z_{12}$ , we can associate a representative point in the complex plane. We see that  $\operatorname{Im}(Z_{12}) = 0$  for a linearly polarized light. Circularly polarized states verify  $\operatorname{Re}(Z_{12}) = 0$  and  $|\operatorname{Im}(Z_{12})| = 1$  and the rest of the points of the plane represent elliptic polarization states. If we now place the condition of orthogonality of two Jones vectors, Eq. (3.1.47), into Eq. (3.1.50), we find that

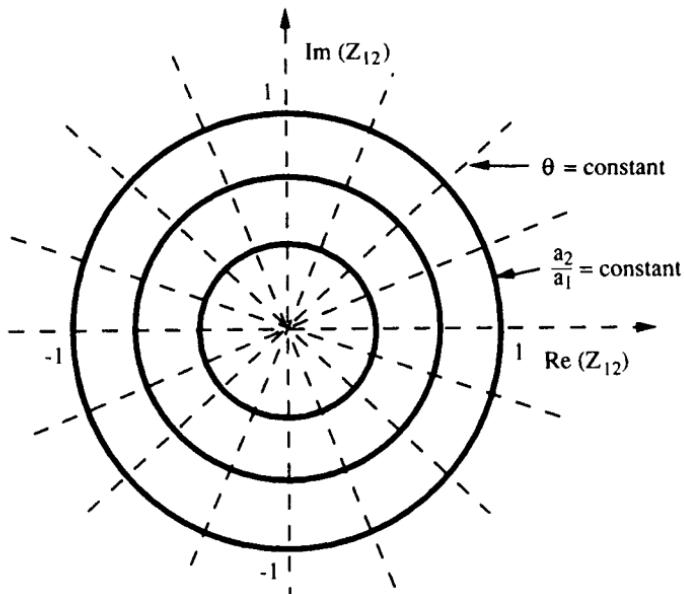
$$Z_{12}^{(a)}(Z_{12}^{(b)})^* = (Z_{12}^{(a)})^*Z_{12}^{(b)} = -1 \quad (3.1.54)$$

for instance, the linear horizontal and vertical states.

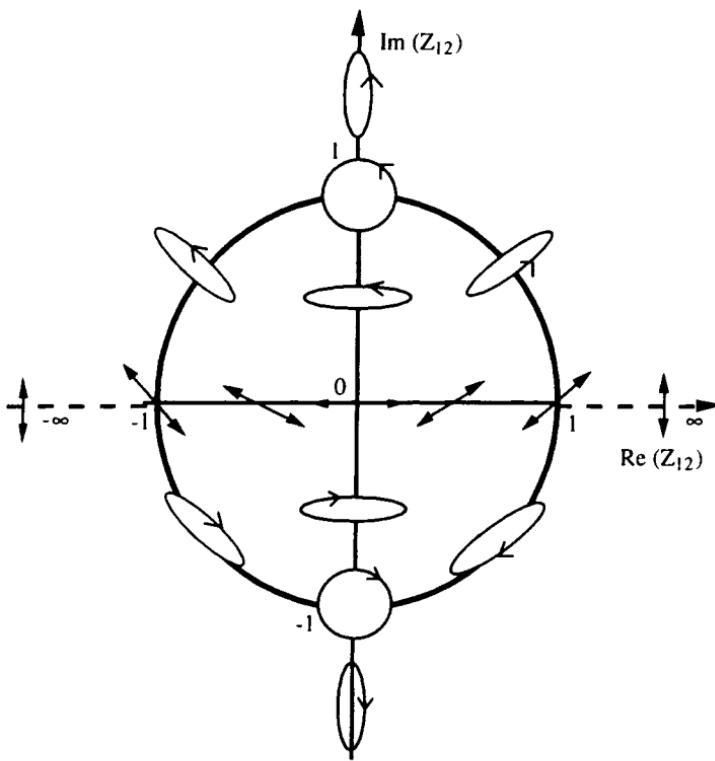
We list below some of the most important properties of  $Z_{12}$ . All the polarization states corresponding to the complex polarization ratio  $Z_{12}$  with the same relative phase angle  $\theta$  lie on a straight line through the origin. On the other hand, polarization states with constant amplitude ratio  $a_2/a_1$  are concentric circles around the origin. The families of curves corresponding to these conditions are orthogonal to each other. A simple geometric construction displayed in Fig. 3.1.4 helps in visualizing these remarkable results.

By way of a summary, we have represented different polarization states in terms of complex mapping with reference to Fig. 3.1.5. Each point on the horizontal axis represents a linear polarized state. The origin corresponds to the linear horizontal state. We see also that for orthogonal states, the major axes of the polarization ellipses are mutually orthogonal and the ellipticities are of the same magnitude with opposite signs. Right-handed (resp. left-handed) elliptical polarization states are localized in the lower (resp. upper) half of the complex plane. Circular polarization states are localized by the intersections of the unit circle with the vertical axis.

Note that Eq. (3.1.50) can be generalized to specify the parameters of the ellipse of polarization in the representation whose basis states are arbitrarily chosen. Let  $Z_{uv}$  represent the generalized complex number in the basis  $(\mathbf{e}_u, \mathbf{e}_v)$ . As the reader can verify, it is a simple matter to prove that the expressions for the azimuth  $\psi$  and the ellipticity  $\chi$  angles, defined in the helicity basis,



**Figure 3.1.4.** Schematic illustration of the condition of constant phase angle  $\theta$  (family of circles) and constant amplitude ratio  $a_2/a_1$  (family of lines orthogonal to those of the first family) in the complex plane.



**Figure 3.1.5.** Correspondence of each point of the complex plane with a unique polarization state. The point at infinity ( $|Z_{12}| = \infty$ ) of the complex plane represents the linear vertical polarization state.

become

$$\psi = \frac{1}{2} \arg(Z_{lr}) \quad (3.1.55)$$

and

$$\tan(2\chi) = \frac{|Z_{lr}| - 1}{|Z_{lr}| + 1} \quad (3.1.56)$$

The values of the complex polarization ratio of some states of polarization expressed in the linear and helicity basis are compared in Table 3.1.2.

It is worth observing that the different complex plane representations are interrelated by bilinear transformations. For example, in the helicity basis, we have

$$Z_{lr} = \frac{-Z_{12} + i}{Z_{12} + i} \quad (3.1.57)$$

**TABLE 3.1.2. Representation of Each Elliptic State of Polarization of Azimuth  $\psi$  and Ellipticity Angle  $\chi$  by a Complex Polarization Ratio  $Z_{12}$  (resp.  $Z_{lr}$ ) in the Linear (resp. helicity) Basis**

Polarization	$\chi$	$\psi$	$Z_{12}$	$Z_{lr}$
Linear horizontal	0	0	0	1
Linear vertical	0	$\frac{\pi}{2}$	$\infty$	-1
Linear +45°	0	$\frac{\pi}{4}$	1	$i$
Linear -45°	0	$-\frac{\pi}{4}$	-1	$-i$
Right-handed circular	$\frac{\pi}{4}$		$-i$	$\infty$
Left-handed circular	$-\frac{\pi}{4}$		$i$	0

To see another illustration of this parametrization, let us note that the normalized Jones vector introduced above can be written in terms of the polarization ratio. We then obtain, if we disregard an unessential absolute phase,

$$\mathbf{E} = \frac{1}{\sqrt{1 + |Z_{uv}|^2}} \left[ \begin{array}{c} 1 \\ Z_{uv} \end{array} \right] \quad (3.1.58)$$

Practical applications of these equations are numerous. We defer a full discussion of the complex polarization ratio formalism to the problem of wave propagation in linear anisotropic media to Section 5.1. However, it is important to appreciate the difficulty in extending these considerations to the case of partially polarized light (see Section 3.3.2).

### 3.1.5. STATISTICAL DESCRIPTION OF FLUCTUATIONS OF A PARTIALLY POLARIZED OPTICAL FIELD

Under the assumption that the wave forms a narrow band about a large central frequency, a single realization of the stochastic analytic signal does not possess a relevant physical meaning, and many of the significant problems reduce

ultimately to the determination of the correlation of the field at different space-time points. This originates from the fact that the electric vector  $\mathbf{E}^{(r)}(\mathbf{R}, t)$  at a point specified by position vector  $\mathbf{R}$  and at time  $t$  is a rapidly fluctuating function of time and the fluctuations at any two points generally will not be the same. We assume that the fluctuations may be described by an ensemble that is stationary, at least in the wide sense ergodic, and have zero mean. We shall now summarize a number of basic properties for the purpose of describing the statistical description of the fluctuations of a partially polarized optical field.

### 3.1.5.1. Correlation Functions of Analytic Signals

We first sketch the proof of the important theorem concerning the cross-correlation of two analytic signals. Let  $F^{(r)}(t)$  and  $G^{(r)}(t)$  be two real stationary fields at a particular point in space and consider the cross-correlation function defined as

$$\langle F(t + \tau)G(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T F(t + \tau; T)G(t; T) dt \quad (3.1.59)$$

where  $F(t; T)$  and  $G(t; T)$  are the analytic signals associated with the truncated functions  $F^{(r)}(t; T)$  and  $G^{(r)}(t; T)$ . We wish to show that  $\langle F(t + \tau)G(t) \rangle = 0$ . The proof goes as follows. First we develop the functions  $F$  and  $G$  into Fourier integrals

$$F(t; T) = \int_0^\infty \hat{F}(v; T) \exp(2i\pi vt) dv \quad (3.1.60a)$$

and

$$G(t; T) = \int_0^\infty \hat{G}(v; T) \exp(2i\pi vt) dv \quad (3.1.60b)$$

Substituting this into the general form Eq. (3.1.59), one finds that

$$\begin{aligned} \langle F(t + \tau)G(t) \rangle &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T dt \int_0^\infty \hat{F}(v; T) \int_0^\infty \hat{G}(v'; T) \\ &\quad \times \exp(2i\pi v\tau) \exp(2i\pi(v + v')t) dv' dv \end{aligned} \quad (3.1.61)$$

Note that  $\lim_{T \rightarrow \infty}$  commutes with the different integrations because  $T$  and  $t$  are independent variables. After interchanging the order of integration and inte-

grating over time, we have

$$\langle F(t + \tau)G(t) \rangle = \lim_{T \rightarrow \infty} \int_0^\infty \int_0^\infty \hat{F}(v; T)\hat{G}(v'; T) \exp(2i\pi v\tau) \frac{\sin\left(\pi(v + v')\frac{T}{2}\right)}{\pi(v + v')\frac{T}{2}} dv dv' \quad (3.1.62)$$

$$= \lim_{T \rightarrow \infty} \int_0^\infty \int_0^\infty \hat{F}(v; T)\hat{G}(v'' - v; T) \times \exp(2i\pi v\tau) \operatorname{sinc}\left((\pi v'')\frac{T}{2}\right) dv'' dv \quad (3.1.63)$$

As  $T \rightarrow \infty$ , the term  $\operatorname{sinc}(\pi(v'')T/2)$  makes the integrand vanish, unless  $v'' = 0$ . But, in that case,  $\hat{F}(v; T)\hat{G}(-v; T)$  vanishes since the spectra  $\hat{F}(v; T)$  and  $\hat{G}(-v; T)$  are zero for negative or zero frequencies. Now, recalling the property of the Dirac delta function

$$\delta(v + v') = \lim_{T \rightarrow \infty} \int_{-T}^T dt \exp(-2i\pi(v + v')t) \quad (3.1.64)$$

it follows that

$$\langle F(t + \tau)G(t) \rangle \equiv 0 \quad (3.1.65)$$

which establishes the theorem.

A special case of the preceding theorem is of particular interest. First let  $F$  and  $G$  be decomposed into real and imaginary parts, respectively:

$$F \equiv F^{(r)} + iF^{(i)} \quad \text{and} \quad G \equiv G^{(r)} + iG^{(i)} \quad (3.1.66)$$

Now substituting Eq. (3.1.66) into Eq. (3.1.65) and on equating real and imaginary parts, we obtain

$$\langle F^{(r)}(t + \tau)G^{(r)}(t) \rangle = \langle F^{(i)}(t + \tau)G^{(i)}(t) \rangle \quad (3.1.67a)$$

$$\langle F^{(i)}(t + \tau)G^{(r)}(t) \rangle = -\langle F^{(r)}(t + \tau)G^{(i)}(t) \rangle \quad (3.1.67b)$$

An additional theorem that concerns the complex cross-correlation is

$$\langle F(t + \tau)G^*(t) \rangle = 2\langle F^{(r)}(t + \tau)G^{(r)}(t) \rangle + 2i\langle F^{(i)}(t + \tau)G^{(r)}(t) \rangle \quad (3.1.68)$$

If we assumed that  $F = G = E_j$  and  $\tau = 0$ , then Eqs. (3.1.67a, b) imply that

$$\langle (E_j^{(r)}(t))^2 \rangle = \langle (E_j^{(i)}(t))^2 \rangle = \frac{1}{2} \langle (E_j(t))^2 \rangle \quad (j = 1, 2) \quad (3.1.69a)$$

$$\langle E_1^{(r)}(t)E_2^{(r)}(t) \rangle = \langle E_1^{(i)}(t)E_2^{(i)}(t) \rangle \quad (3.1.69b)$$

$$\langle E_1^{(r)}(t)E_2^{(i)}(t) \rangle = -\langle E_1^{(i)}(t)E_2^{(r)}(t) \rangle \quad (3.1.69c)$$

$$\langle E_1^{(r)}(t)E_1^{(i)}(t) \rangle = \langle E_2^{(i)}(t)E_2^{(r)}(t) \rangle = 0 \quad (3.1.69d)$$

Thus at the same instant  $t$ ,  $E_j^{(r)}$  and  $E_j^{(i)}$  are uncorrelated and have identical variances. In summary, we indicate the following important relations involving correlation functions of analytic signals:

$$\langle (E_1)^2 \rangle = 2\langle (E_1^{(r)})^2 \rangle = 2\langle (E_1^{(i)})^2 \rangle \quad (3.1.70a)$$

$$\langle E_1 E_2^* \rangle = 2\langle E_1^{(r)} E_2^{(r)} \rangle - 2i\langle E_1^{(i)} E_2^{(i)} \rangle \quad (3.1.70b)$$

Summarizing, we have shown that the real and imaginary parts of the transverse components of the fluctuating electric field vector are pairwise-independent and the real and the imaginary parts of the same component are independent with the same variance.

### 3.1.5.2. The Probability Density Function of the Radiation Field

The most complete specification of the random field  $\mathbf{E}$  requires the knowledge of the hierarchy of  $N$ -fold joint probability density functions (PDF) for all  $N$ :

$$p_N(\mathbf{E}_1(\mathbf{R}_1, t_1; T), \mathbf{E}_2(\mathbf{R}_2, t_2; T), \dots, \mathbf{E}_j(\mathbf{R}_j, t_j; T), \dots, \mathbf{E}_N(\mathbf{R}_N, t_N; T)) \quad (3.1.71)$$

such that

$$p_N(\mathbf{E}_1(\mathbf{R}_1, t_1; T), \mathbf{E}_2(\mathbf{R}_2, t_2; T), \dots, \mathbf{E}_j(\mathbf{R}_j, t_j; T), \dots, \mathbf{E}_N(\mathbf{R}_N, t_N; T)) \prod_{j=1}^N d^2 \mathbf{E}_j(\mathbf{R}_j, t_j; T) \quad (3.1.72)$$

represents the probability of finding a value between  $\mathbf{E}_j$  and  $\mathbf{E}_j + d^2 \mathbf{E}_j$ ,  $j = 1, N$  [where the measure has been defined such  $d^2 \mathbf{E}_j \equiv d\mathbf{E}_j^{(r)} + id\mathbf{E}_j^{(i)}$ ] for the field variable at space-time points  $(\mathbf{R}_j, t_j)$ . Once the distribution  $p_N$  is known, we can compute any observable function  $f$  of the field at  $N$  space-time points by

evaluating the ensemble average:

$$\langle f(E_1, E_2, \dots, E_j, \dots, E_N) \rangle$$

$$= \int_{N - \text{fold integral}} \cdots \int f(E_1, E_2, \dots, E_j, \dots, E_N) p_N(E_1, E_2, \dots, E_j, \dots, E_N) \prod_{k=1}^N d^2 E_k$$

Instead of using PDF to describe the ensemble, it may be characterized by its moments. For this purpose, we set  $f(E_1, E_2, \dots, E_j, \dots, E_N) = \prod_{k=1}^N E_k^{s_k}$  in Eq. (3.1.73). The ensemble average function  $\langle f(E_1(R_1, t_1), E_2(R_2, t_2), \dots, E_j(R_j, t_j), \dots, E_N(R_N, t_N)) \rangle$  is, in general, time-dependent. However, if  $p_N$  and consequently all moments of  $\mathbf{E}$  are invariant relative to the translation over the entire set of times  $t_j$ , the process is termed *stationary*. If  $p_N$  and consequently all moments of  $\mathbf{E}$  are invariant relative to the translation over the entire set of position vectors  $\mathbf{R}_j$ , the process is termed *spatially homogeneous* (i.e., all points in space are statistically equivalent).

The sharp brackets denote an ensemble average over an ensemble of different realizations of the field or a time average of a typical member of the ensemble of realizations because of our assumption of stationarity and ergodicity. As the field is taken as homogeneous and stationary, the PDF  $p_N$  is invariant relative to the translation of the entire set of points  $\mathbf{R}_1, t_1, \mathbf{R}_2, t_2, \dots, \mathbf{R}_N, t_N$  in space and in time.

As a practical matter, only a small finite number of moments of the electric field may be determined directly either experimentally or theoretically. For example, a complete specification of radiation fields obeying Gaussian statistics implies knowledge of second-order moments only. However, we should not end this discussion without pointing out that there are practical cases where it is appropriate to use higher-order correlation functions to describe the coherence properties of the radiation field, such as light from nonthermal sources. We conclude this section with a brief look at the higher-order coherence properties of the field. The precise definition of the different orders of coherence is in terms of the spectral density, which is a many-body function. One constructs the  $(n + m)$ th-order correlations functions of the field as the ensemble average

$$\Phi_{s_1 \dots s_{n+m}}^{n,m}(\mathbf{R}_1, t_1; \dots; \mathbf{R}_{n+m}, t_{n+m}) \equiv \left\langle \prod_{j=1}^n E_{s_j}^*(\mathbf{R}_j, t_j) \prod_{k=n+1}^{n+m} E_{s_k}(\mathbf{R}_k, t_k) \right\rangle \quad (3.1.74)$$

where the  $s_j$  terms are polarization indices that label the Cartesian components of the complex analytic signal  $\mathbf{E}$ . These moments are, in general, difficult to compute, but they have the following properties. When the field is statistically stationary,  $\Phi_{s_1 \dots s_{n+m}}^{n,m}(\mathbf{R}_1, t_1; \dots; \mathbf{R}_{n+m}, t_{n+m})$  is invariant with respect to translation of the origin of time and depends then only on  $n + m - 1$  time arguments, that is, on the differences  $\tau_k \equiv t_k - t_1$ ,  $k = 2, \dots, n + m$ . To pursue this point further, the interested reader is invited to consult Refs. 1, 2, 8, and 10.

### 3.1.5.3. Second-Order Coherence Properties of a Stochastic Radiation Field

In our earlier discussion, it was pointed out that the full description by the  $N$ -fold joint PDF  $p_N$  is seldom possible and in most practical applications only a first-order description (i.e., at a single instant) or a second-order description (i.e., at two time instants) is possible. From this point on, we develop the study of the twofold statistical properties of light. The correlation between the electric vectors  $\mathbf{E}(\mathbf{R}, t; T)$  at any two points constitutes the phenomenon of partial coherence, while the correlation between the orthogonal components of  $\mathbf{E}(\mathbf{R}, t; T)$  at any one point constitutes the phenomenon of partial polarization.

It will be useful to begin by recalling some standard results relating to partial coherence theory [1, 8, 10]. By virtue of the stationary property of the radiation, the correlation that exists between the electric field vectors at two points (specified by position vectors  $\mathbf{R}_1$  and  $\mathbf{R}_2$ ) at time instants separated by a time interval  $\tau$  may be defined as

$$\begin{aligned}\Phi_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau) &= \langle E_j(\mathbf{R}_1, t + \tau) E_k^*(\mathbf{R}_2, t) \rangle \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E_j(\mathbf{R}_1, t + \tau; T) E_k^*(\mathbf{R}_2, t; T) dt\end{aligned}\quad (3.1.75)$$

Using the ergodic assumption, the ensemble average may be replaced by the corresponding temporal average. Then  $\langle E_j(\mathbf{R}_1, t + \tau) E_k^*(\mathbf{R}_2, t) \rangle$  form the components of the mutual coherence (also termed *cross-covariance*) tensor  $\Phi_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau)$ <sup>5</sup>. The simple symmetry property

$$\Phi_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau) = \Phi_{kj}^*(\mathbf{R}_2, \mathbf{R}_1, -\tau) \quad (3.1.76)$$

follows from the assumption of stationary of the fluctuating field and Eq. (3.1.75).

Similarly, when  $\mathbf{R}_1 = \mathbf{R}_2$ , the  $\Phi_{jk}(\mathbf{R}, \mathbf{R}, \tau)$  defines the autocorrelation tensor of  $\mathbf{E}(\mathbf{R}, t)$ . Here, for brevity,  $\Phi(\tau)$  is written in place of  $\Phi(\mathbf{R}, \mathbf{R}, \tau)$ . We can summarize the general properties of the correlation tensors and discuss the physical consequences. It may be shown that the  $\Phi_{jk}$  values fall off sufficiently

<sup>5</sup>Note that  $\Phi_2$  is itself a complex analytic signal and can be separated into real and imaginary parts (like in Eq. (3.1.8))

$$\Phi_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau) = \Phi_{jk}^{(r)}(\mathbf{R}_1, \mathbf{R}_2, \tau) + i\Phi_{jk}^{(i)}(\mathbf{R}_1, \mathbf{R}_2, \tau)$$

where  $\Phi_{jk}^{(r)}$  and  $\Phi_{jk}^{(i)}$  form a pair of conjugate functions, that is,

$$\Phi_{jk}^{(r)}(\mathbf{R}_1, \mathbf{R}_2, \tau) = -\frac{1}{\pi} \text{PP} \left( \int_{-\infty}^{\infty} \Phi_{jk}^{(i)} \left( \frac{\mathbf{R}_1, \mathbf{R}_2, \tau'}{\tau' - \tau} \right) d\tau' \right).$$

Similar results hold for  $\Xi$  and  $\Pi$ .

rapidly to zero as  $|\tau| \rightarrow \infty$  to ensure that they are absolutely integrable with respect to  $t$  so that they possess a Fourier transform, and

$$\Phi_{jj}(-\tau) = \Phi_{jj}(\tau) \quad (3.1.77a)$$

$$\Phi_{jk}(-\tau) = \Phi_{kj}^*(\tau) \quad (3.1.77b)$$

and

$$|\Phi_{jk}(\tau)|^2 \leq \Phi_{jj}(0)\Phi_{kk}(0) \quad (3.1.77c)$$

Equations (3.1.77a, b) are a consequence of homogeneity; Eq. (3.1.77c) follows from the Cauchy-Schwarz inequality.

For a stationary field, these second moments depend on  $t + \tau$  and  $t$  only through the time difference  $\tau$  and are assumed to be continuous at  $\tau = 0$ . One further remark about notation: temporal coherence is characterized by  $\Phi(\mathbf{R}, \mathbf{R}, \tau)$  where the two points  $\mathbf{R}_1$  and  $\mathbf{R}_2$  coincide and spatial coherence is characterized by  $\Phi(\mathbf{R}_1, \mathbf{R}_2, 0)$  where the time delay is kept fixed.

Let us also introduce formally the normalized complex electric correlation tensor by setting

$$\gamma_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau) \equiv \frac{\Phi_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau)}{[\Phi_{jj}(\mathbf{R}_1, \mathbf{R}_2, \tau)\Phi_{kk}(\mathbf{R}_1, \mathbf{R}_2, \tau)]^{1/2}} \quad (3.1.78)$$

By the Cauchy-Schwarz inequality, we can prove that  $0 \leq |\gamma_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau)| \leq 1$ . We also define a complex correlation factor

$$\begin{aligned} \gamma(\mathbf{R}, \mathbf{R}, 0) \equiv \gamma_{12}(0) &= \frac{\Phi_{12}(\mathbf{R}, \mathbf{R}, 0)}{[\Phi_{11}(\mathbf{R}, \mathbf{R}, 0)\Phi_{22}(\mathbf{R}, \mathbf{R}, 0)]^{1/2}} \\ &= \frac{\langle E_1(\mathbf{R}, t + \tau)E_2^*(\mathbf{R}, t) \rangle}{(\langle |E_1(\mathbf{R}, t)|^2 \rangle)^{1/2}(\langle |E_2(\mathbf{R}, t)|^2 \rangle)^{1/2}} \end{aligned} \quad (3.1.79)$$

This definition is of particular importance in the description of interference fringes in two-beam interferometry:  $\gamma$  is a measure of the visibility of the interference fringes at point characterized by position vector  $\mathbf{R}$  [1-2, 8, 10].

In a manner identical to that used in arriving at Eq. (3.1.75), one may define second-order correlation tensors, assuming again that the fields are stationary and ergodic, which characterize the correlation that exists between the magnetic field vectors (or the electric vector and the magnetic vector) at two typical points ( $\mathbf{R}_1$  and  $\mathbf{R}_2$ ) in space at time instants separated by a time interval  $\tau$  by

$$\begin{aligned} \Xi_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau) &= \langle H_j(\mathbf{R}_1, t + \tau)H_k^*(\mathbf{R}_2, t) \rangle \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T H_j(\mathbf{R}_1, t + \tau; T)H_k^*(\mathbf{R}_2, t; T) dt \end{aligned} \quad (3.1.80)$$

and

$$\begin{aligned}\Pi_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau) &= \langle E_j(\mathbf{R}_1, t + \tau) H_k^*(\mathbf{R}_2, t) \rangle \\ &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E_j(\mathbf{R}_1, t + \tau; T) H_k^*(\mathbf{R}_2, t; T) dt\end{aligned}\quad (3.1.81)$$

where, as before,  $\mathbf{E}(\mathbf{R}, t)$  and  $\mathbf{H}(\mathbf{R}, t)$  denote the complex analytic signal representations associated with the real electric field  $E^{(r)}(\mathbf{R}, t)$  and the real magnetic field  $H^{(r)}(\mathbf{R}, t)$  at a point specified by a position vector  $\mathbf{R}$  at time  $t$ , respectively.

### 3.1.5.4. Relations Between Correlation Tensors and Poynting Vector

To establish the connection between the Poynting vector  $\mathbf{P}$  and the coherency matrix elements, we proceed as follows. The average electric energy density can be expressed in terms of

$$w_e \equiv \langle (\mathbf{E}^{(r)}(\mathbf{R}, t))^2 \rangle \quad (3.1.82)$$

Making use of Eq. (3.1.8), we obtain the following simple expression:

$$\langle \mathbf{E} \cdot \mathbf{E}^* \rangle = \langle (\mathbf{E}^{(r)}(\mathbf{R}, t))^2 \rangle + \langle (\mathbf{E}^{(i)}(\mathbf{R}, t))^2 \rangle + 2i \langle \mathbf{E}^{(r)}(\mathbf{R}, t) \cdot \mathbf{E}^{(i)}(\mathbf{R}, t) \rangle \quad (3.1.83)$$

On substitution from Eqs. (3.1.66a, d) into Eqs. (3.1.59), we find the relation between the average electric energy density of a stationary radiation field in free space and the electric coherency matrix

$$w_e = \frac{1}{2} \langle \mathbf{E} \cdot \mathbf{E}^* \rangle = \text{tr}(\Phi(\mathbf{R}, \mathbf{R}, 0)) \quad (3.1.84)$$

As usual, the factor of  $\frac{1}{2}$  reflects the fact that  $E_j^{(r)}$  and  $E_j^{(i)}$  have identical variances (Eq. (3.1.69a)). Similarly, one finds that Eq. (3.1.80) simply asserts that the magnetic energy density

$$w_m = \frac{1}{2} \langle \mathbf{H} \cdot \mathbf{H}^* \rangle = \frac{1}{2} \text{tr}(\Xi(\mathbf{R}, \mathbf{R}, 0)) \quad (3.1.85)$$

We have seen previously (Section 2.2.3) that the average Poynting vector  $\mathbf{P}$  of an electromagnetic field writes

$$\langle \mathbf{P} \rangle = \langle \mathbf{E}^{(r)}(\mathbf{R}, t) \times \mathbf{H}^{(r)}(\mathbf{R}, t) \rangle = \frac{1}{2} \text{Re} \langle \mathbf{E}(\mathbf{R}, t) \times \mathbf{H}^*(\mathbf{R}, t) \rangle \quad (3.1.86)$$

We have at once, on making use of Eq. (2.3.9), that

$$\langle \mathbf{P} \rangle = \frac{1}{4\pi\nu\mu\mu_0} \operatorname{Re} \langle \mathbf{E} \times (\mathbf{q}^* \times \mathbf{E}^*) \rangle \quad (3.1.87)$$

On the right-hand side of Eq. (3.1.87), we make use of the triple cross-product vector identity  $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c})\mathbf{b} - (\mathbf{a} \cdot \mathbf{b})\mathbf{c}$ , where  $\mathbf{a}$ ,  $\mathbf{b}$ , and  $\mathbf{c}$  are three arbitrary vectors. If we substitute Eq. (3.1.84) into Eq. (3.1.87), we find, for a transverse wave propagating in the direction characterized by unit vector  $\mathbf{e}_3$ , that

$$\langle \mathbf{P} \rangle = \frac{1}{2} \left( \frac{\varepsilon\varepsilon_0}{\mu\mu_0} \right)^{1/2} \langle \mathbf{E} \cdot \mathbf{E}^* \rangle \mathbf{e}_3 = \left( \frac{\varepsilon\varepsilon_0}{\mu\mu_0} \right)^{1/2} w_e \mathbf{e}_3 = \left( \frac{\varepsilon\varepsilon_0}{\mu\mu_0} \right)^{1/2} \operatorname{tr}(\Phi(\mathbf{R}, \mathbf{R}, 0)) \mathbf{e}_3 \quad (3.1.88)$$

This expression looks almost like Eq. (2.3.15) and may be regarded as a generalization to a quasimonochromatic wavefield of the result obtained for a strictly monochromatic wavefield in Part 2.

We have introduced the mixed coherency matrix  $\Pi_{jk}(\mathbf{R}_1, \mathbf{R}_2, t)$  via Eq. (3.1.81). Using this notation, Eq. (3.1.86) can be recast into the form

$$\langle \mathbf{P}_i \rangle = \varepsilon_{ijk} \operatorname{Re} \langle E_j H_k^* - E_k H_j^* \rangle = \varepsilon_{ijk} \operatorname{Re} (\Pi_{jk}(\mathbf{R}, \mathbf{R}, 0) - \Pi_{jk}^*(\mathbf{R}, \mathbf{R}, 0)) \quad (3.1.89)$$

where  $\varepsilon_{ijk}$  is the completely antisymmetric unit tensor of Levi-Cevita, that is,  $\varepsilon_{ijk} = 1$  or  $-1$  according to whether  $(i, j, k)$  is an even or odd permutation of  $(1, 2, 3)$  and  $\varepsilon_{ijk} = 0$  when two or more of the subscripts are equal.

### 3.1.5.5. Dynamical Equations for Second-Order Correlation Tensors

As was done for the propagation of fields in Part 2, we can ask for the differential equations for the correlation tensors that describe the evolution of their spatial and temporal characteristics on propagation. We now outline a derivation of the dynamical equations governing the propagation for these tensors originally due to Wolf [1]. Following the formalism developed in Section 3.1.2, it is easy to show that the analytic signals  $E_j(\mathbf{R}, t)$  associated with sample functions of the real Cartesian components  $E_j^{(r)}(\mathbf{R}, t)$  that characterize the fluctuating optical transverse field at the point  $\mathbf{R}$  and at time  $t$ , obey in free space, of wave equation

$$\nabla^2 \mathbf{E}(\mathbf{R}, t) = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}(\mathbf{R}, t)}{\partial t^2} \quad (3.1.90)$$

Consequently, we obtain outside the domain occupied by the source

$$\nabla_1^2 \mathbf{E}_j(\mathbf{R}_1, t_1) = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}_j(\mathbf{R}_1, t_1)}{\partial t_1^2} \quad (j = 1, 2) \quad (3.1.91)$$

by setting  $\mathbf{R} = \mathbf{R}_1$  and  $t = t_1$ . The symbol  $\nabla_1^2 \equiv \partial^2/\partial x_1^2 + \partial^2/\partial y_1^2 + \partial^2/\partial z_1^2$

denotes the Laplacian operating on the coordinates of the point  $\mathbf{R}_1$  of coordinates  $(x_1, y_1, z_1)$ . Next multiply both sides of Eq. (3.1.91) by  $E_k^*(\mathbf{R}_2, t_2)$ . We have

$$\nabla_1^2(E_j(\mathbf{R}_1, t_1)E_k^*(\mathbf{R}_2, t_2)) = \frac{1}{c^2} \frac{\partial^2(E_j(\mathbf{R}_1, t_1)E_k^*(\mathbf{R}_2, t_2))}{\partial t_1^2} \quad (3.1.92)$$

The ensemble average over the different realizations of the field permits us to rewrite Eq. (3.1.92) in the final form

$$\nabla_1^2\Phi_{jk}(\mathbf{R}_1, t_1; \mathbf{R}_2, t_2) = \frac{1}{c^2} \frac{\partial^2\Phi_{jk}(\mathbf{R}_1, t_1; \mathbf{R}_2, t_2)}{\partial t_1^2} \quad (3.1.93)$$

In a strictly similar way, we have the analogous relation

$$\nabla_2^2\Phi_{jk}(\mathbf{R}_1, t_1; \mathbf{R}_2, t_2) = \frac{1}{c^2} \frac{\partial^2\Phi_{jk}(\mathbf{R}_1, t_1; \mathbf{R}_2, t_2)}{\partial t_2^2} \quad (3.1.94)$$

where  $\nabla_2^2$  indicates the Laplacian operating on the coordinates of the point  $R_2$ . If the process is stationary in time, the mutual coherence tensor  $\Phi(\mathbf{R}_1, t_1; \mathbf{R}_2, t_2)$  becomes  $\Phi(\mathbf{R}_1, \mathbf{R}_2, t_1 - t_2 = \tau)$  and the second-order correlation tensor  $\Phi$  obeys the propagation equations in vacuo:

$$\nabla_l^2\Phi(\mathbf{R}_1, \mathbf{R}_2, \tau) = \frac{1}{c^2} \frac{\partial^2\Phi(\mathbf{R}_1, \mathbf{R}_2, \tau)}{\partial \tau^2} \quad (l = 1, 2) \quad (3.1.95)$$

As a result, the correlation tensor  $\Phi$  obeys two wave equations, one for each of the space variables. Consequently, the correlation tensor  $\Phi$  is required to obey the same propagation equation and boundary conditions as the electric field vector, which we introduced in Section 3.2. The usefulness of Eq. (3.1.95) is that it couples the dependence of  $\Phi(\mathbf{R}_1, \mathbf{R}_2, \tau)$  on spatial variables  $\mathbf{R}_1$  and  $\mathbf{R}_2$  (left-hand side) to the dependence of  $\Phi(\mathbf{R}_1, \mathbf{R}_2, \tau)$  on  $\tau$  (right-hand side). In view of this remark, we can expect that spatial and temporal coherence properties of the radiation field are not independent to each other. Similar equations may be derived easily for the correlation tensors  $\Pi$  and  $\Xi$ . Specifically, these dynamical equations describe the changes in the correlation tensors as one of the two points ( $\mathbf{R}_1$  or  $\mathbf{R}_2$ ) is kept fixed while the other one and  $\tau$  vary. The importance of Eq. (3.1.95) has been discussed in the literature [1, 2, 8, 10].

### 3.1.5.6. Wiener-Khintchine Theorem and Cross-Spectral Correlation Tensor

It is convenient, as starting point, to form the product  $\hat{E}_j(\mathbf{R}_1, v + v_1)\hat{E}_k^*(\mathbf{R}_2, v)$  and ensemble average. After interchanging the order of integration and

averaging, we obtain

$$\langle \hat{E}_j(\mathbf{R}_1, v + v_1) \hat{E}_k^*(\mathbf{R}_2, v) \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle E_j(\mathbf{R}_1, t_1) E_k^*(\mathbf{R}_2, t_2) \rangle \\ \times \exp(2i\pi v(t_1 - t_2)) \exp(2i\pi v_1 t_1) dt_1 dt_2 \quad (3.1.96a)$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle E_j(\mathbf{R}_1, t + \tau) E_k^*(\mathbf{R}_2, t) \rangle \\ \times \exp(2i\pi(v + v_1)\tau) \exp(2i\pi v_1 t) dt d\tau \quad (3.1.96b)$$

where we have set  $t_1 = t + \tau$  and  $t_2 = t$ . If we substitute from Eq. (3.1.96b) into Eq. (3.1.75), and carry out the trivial integration with respect to  $t$ , we obtain

$$\langle \hat{E}_j(\mathbf{R}_1, v + v_1) \hat{E}_k^*(\mathbf{R}_2, v) \rangle = \delta(v_1) \hat{\Phi}_{jk}(\mathbf{R}_1, \mathbf{R}_2, v) \quad (3.1.97)$$

where

$$\hat{\Phi}_{jk}(\mathbf{R}_1, \mathbf{R}_2, v) = \int_{-\infty}^{\infty} \Phi_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau) \exp(2\pi i v \tau) d\tau \quad (3.1.98)$$

are the elements of the cross-spectral tensor of the electric field at points  $\mathbf{R}_1$  and  $\mathbf{R}_2$ . Equation (3.1.97) is important because it shows that the components of the cross-spectral tensors belonging to different frequencies are uncorrelated. Equation (3.1.98) is termed the *generalized Wiener–Khintchine theorem*. As the name implies, the theorem was first demonstrated in papers by Wiener [11] and Khintchine [12]. The Wiener–Khintchine theorem, stated mathematically in Eq. (3.1.98), can be expressed in words as follows. The cross-spectral density tensor  $\hat{\Phi}(\mathbf{R}_1, \mathbf{R}_2, v)$  of the fluctuating optical field can be derived from a Fourier transform of the corresponding cross-correlation tensor  $\Phi(\mathbf{R}_1, \mathbf{R}_2, \tau)$ , so that knowledge of either one implies the other.

From the quasimonochromatic approximation, that is, where the integral will vanish for any  $v'$  that is much outside the range  $v' \ll \Delta v$ , we can rewrite Eq. (3.1.50) as

$$\Phi_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau) = \int_0^{\infty} \hat{\Phi}_{jk}(\mathbf{R}_1, \mathbf{R}_2, v) \exp(-2\pi i v \tau) dv \quad (3.1.99)$$

$$= \exp(-2\pi i v_0 \tau) \int_0^{\infty} \hat{\Phi}_{jk}(\mathbf{R}_1, \mathbf{R}_2, v_0 + v') \exp(-2\pi i v' \tau) dv' \quad (3.1.100)$$

where  $v_0$  is the central frequency of the spectral distribution whose width  $\Delta v \ll v_0$ . In this spectral range, we have  $2\pi i v' \tau \ll 1$  for any  $\tau \ll 1/\Delta v$ , and Eq.

(3.1.100) becomes

$$\Phi_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau) = \exp(-2\pi i v_0 \tau) \Phi_{jk}(\mathbf{R}_1, \mathbf{R}_2, 0) \quad (3.1.101)$$

for  $|\tau| \ll \tau_2$ , specifically, periodic in  $\tau$ . For long time delays, the full descriptions such as given by (3.1.99a) must be employed.

Following the formalism developed in section 3.1.5.5, we can derive the differential equations for the propagation of the cross-spectral correlation tensor. On expressing  $\Phi_{jk}(\mathbf{R}_1, \mathbf{R}_2, \tau)$  in terms of its Fourier transform  $\hat{\Phi}_{jk}(\mathbf{R}_1, \mathbf{R}_2, v)$  and making use of the dynamical equations, Eqs. (3.1.93) and (3.1.94), we can derive the free-space Helmholtz equations for the propagation of the cross-spectral density

$$\nabla_1^2 \hat{\Phi}_{jk}(\mathbf{R}_1, \mathbf{R}_2, v) + \frac{4\pi^2 v^2}{c^2} \hat{\Phi}_{jk}(\mathbf{R}_1, \mathbf{R}_2, v) = 0 \quad (3.1.102a)$$

$$\nabla_2^2 \hat{\Phi}_{jk}(\mathbf{R}_1, \mathbf{R}_2, v) + \frac{4\pi^2 v^2}{c^2} \hat{\Phi}_{jk}(\mathbf{R}_1, \mathbf{R}_2, v) = 0 \quad (3.1.102b)$$

which we have already encountered in Part 2, [Eq. (2.3.6)].

The normalized complex spectral correlation tensor can now be written by following the approach we employed to derive Eq. (3.1.78):

$$\mu_{jk}(\mathbf{R}_1, \mathbf{R}_2, v) \equiv \frac{\hat{\Phi}_{jk}(\mathbf{R}_1, \mathbf{R}_2, v)}{[\hat{\Phi}_{jj}(\mathbf{R}_1, \mathbf{R}_2, v) \hat{\Phi}_{kk}(\mathbf{R}_1, \mathbf{R}_2, v)]^{1/2}} \quad (3.1.103)$$

Unlike the fact that  $\Phi(\mathbf{R}_1, \mathbf{R}_2, v)$  and  $\hat{\Phi}(\mathbf{R}_1, \mathbf{R}_2, v)$  form a Fourier transform pair,  $\gamma(\mathbf{R}_1, \mathbf{R}_2, v)$  and  $\mu(\mathbf{R}_1, \mathbf{R}_2, v)$  are in general, not Fourier transforms of each other. Examination of Eq. (3.1.103) reveals that

$$0 \leq |\mu_{jk}(\mathbf{R}_1, \mathbf{R}_2, v)| \leq 1 \quad (3.1.104)$$

Reference is made to Mandel and Wolf for general properties of the cross-spectral correlation tensor [1].

### 3.1.5.7. Blackbody Radiation in Equilibrium in an Enclosure

As an illustration of the preceding formulas, it is instructive to describe the temporal coherence of blackbody radiation in equilibrium in an enclosure of volume  $V$  and temperature  $T$ . Blackbody radiation is an important model of a radiation source and holds a special place in physics because of its historical role in the development of quantum physics. According to the basic principles of statistical mechanics, a blackbody in local thermodynamic equilibrium with uniform temperature is an object that perfectly absorbs and reradiates all the incident radiation incident on it. It is worth emphasizing that Bourret was the

first to derive an expression of the second-order electric correlation tensor of blackbody radiation based on an important hydrodynamical analogy with the theory of homogeneous and isotropic turbulence of an incompressible fluid [13]. See Appendix A for details.

We consider a radiation emitted into vacuum by a blackbody with uniform temperature  $T$  and whose linear dimensions are large compared to the mean wavelength of the radiation. Hence the field is isotropic and  $\Phi_{ij}$  and  $\gamma_{ij}$  depend on  $\mathbf{R}_1$  and  $\mathbf{R}_2$  through the difference only  $\mathbf{R} = \mathbf{R}_1 - \mathbf{R}_2$ . The theory of Kano and Wolf is an appropriate starting point on which to base a study of temporal coherence [14]. Here we state without proof the expression for the second-order complex electric normalized correlation tensor  $\gamma_{ij}$  of blackbody radiation

$$\gamma_{ij}(\mathbf{R}, \tau) = \frac{45a^4}{8\pi^5} \int \frac{q^2 \delta_{ij} - q_i q_j}{q(\exp(aq) - 1)} \exp(i(\mathbf{q} \cdot \mathbf{R} - q\tau)) d\mathbf{q} \quad (3.1.105)$$

where we have set  $a \equiv \hbar c/kT$ , and the integration in (3.1.105) extends over the whole  $\mathbf{q}$  space. Equation (3.1.105) can be also rewritten as

$$\gamma_{ij}(\mathbf{R}, \tau) = \frac{45a^4}{8\pi^5} \left( \frac{\partial^2}{\partial R_i \partial R_j} - \delta_{ij} \nabla^2 \right) \int \frac{\exp(i(\mathbf{q} \cdot \mathbf{R} - q\tau))}{q(\exp(aq) - 1)} d\mathbf{q} \quad (3.1.106)$$

This integral has a rather complex structure, but it is advantageous to convert Eq. (3.1.106) to polar coordinates. By usual rules of differentiation, Mehta and Wolf found that  $\gamma_{ij}(\mathbf{R}, \tau)$  is given by [15]

$$\begin{aligned} \gamma_{ij}(\mathbf{R}, \tau) &= \frac{45a^4}{8\pi^5} \left( \frac{\partial^2}{\partial R_i \partial R_j} - \delta_{ij} \nabla^2 \right) \int_0^\infty \frac{q \exp(-iqc\tau)}{(\exp(aq) - 1)} dq \\ &\times \int_0^\pi \exp(iqR \cos(\theta)\sin(\theta)) d\theta \int_0^{2\pi} d\phi \\ &= \frac{45a^4}{2\pi^4} \left( \frac{\partial^2}{\partial R_i \partial R_j} - \delta_{ij} \nabla^2 \right) \frac{1}{R} \int_0^\infty \frac{\sin(qR) \exp(-iqc\tau)}{(\exp(aq) - 1)} dq \end{aligned} \quad (3.1.107)$$

If we use

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{1}{2i} \int_0^\infty & (\exp(iq(R - c\tau)) - \exp(-iq(R + c\tau))) \exp(-naq) dq \\ &= \frac{1}{2i} \sum_{n=1}^{\infty} \left( \frac{1}{na - i(R - c\tau)} - \frac{1}{na + i(R + c\tau)} \right) \end{aligned} \quad (3.1.108)$$

$\gamma_{ij}(\mathbf{R}, \tau)$  can be expressed by the formula

$$\gamma_{ij}(\mathbf{R}, \tau) = \frac{45a^4}{2\pi^4} \left( \frac{\partial^2}{\partial R_i \partial R_j} - \delta_{ij} \nabla^2 \right) \sum_{n=1}^{\infty} \frac{1}{(na + i\tau)^2 + R^2} \quad (3.1.109)$$

Finally, if we carry out the differentiation on the right-hand side of Eq. (3.1.109), we have

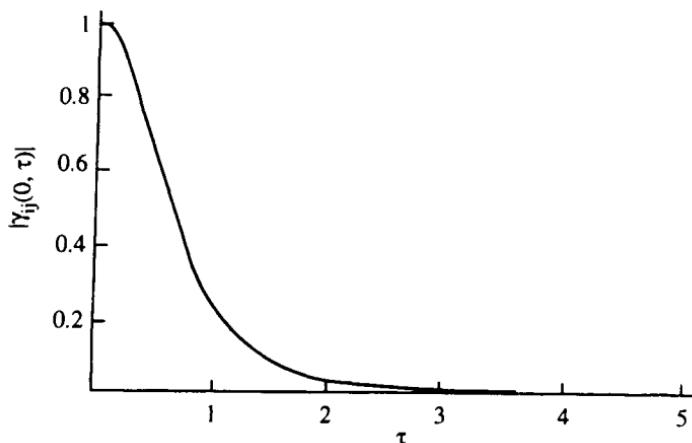
$$\gamma_{ij}(\mathbf{R}, \tau) = \frac{90a^4}{\pi^4} \sum_{n=1}^{\infty} \frac{\delta_{ij}}{((na + i\tau)^2 + R^2)^2} + 2 \frac{R_i R_j - R^2 \delta_{ij}}{((na + i\tau)^2 + R^2)^3} \quad (3.1.110)$$

It is interesting to examine the behavior of the temporal degree of coherence by setting  $\mathbf{R} = 0$ . We then obtain the diagonal Kano-Wolf formula [14]

$$\gamma_{ij}(0, \tau) = \frac{90}{\pi^4} \zeta \left( 4, 1 + \frac{i\tau}{a} \right) \delta_{ij} \quad (3.1.111)$$

where the generalized Riemann-Hurwitz zeta function is given by  $\zeta(s, a) = \sum_{k=0}^{\infty} (k+a)^{-s}$ . Note that  $\zeta(4, 0) = \pi^4/90$ . As a concrete example, Fig. 3.1.6 shows the behavior of the modulus of  $\gamma_{ij}(0, \tau)$  for positive values of time delay  $\tau$ . Time is scaled by  $\tau \rightarrow \tau(kT/\hbar)$ . The corresponding values for negative  $\tau$  are obtained from the parity property of the modulus  $|\gamma_{ij}(0, -\tau)| = |\gamma_{ij}(0, \tau)|$ .

From Fig. 3.1.6, one observes that  $|\gamma_{ij}(0, \tau)|$  decreases monotonically with the time delay  $\tau$  and that the range of correlations in blackbody radiation is  $\sim (\hbar/kT)$  in time. For a more detailed discussion of some of the topics treated in this section, the reader may consult Refs. 14 and 15.



**Figure 3.1.6.** Modulus of  $\gamma_{ij}(0, \tau)$  for blackbody radiation as a function of the dimensionless time  $(kT/\hbar)\tau$ .

### 3.1.6. STOKES PARAMETERS AND COHERENCY MATRIX FORMALISM

In most physical realizable situations, the wave is not in a pure state. Thus the knowledge of the parameters describing the optical field is incomplete. For practical measurements we must average over an ensemble (or over time) of fields. A useful way to handle this averaging is through the use of Stokes parameters and of the coherency or density matrix. This section is devoted to the analysis of these formalisms and their physical consequence.

#### 3.1.6.1. Stokes Parameters

Let us consider a uniform, statistically stationary quasimonochromatic plane wavefield, of central frequency  $\nu_0$  very large compared to the spectral bandwidth  $\Delta\nu$ , propagating in the direction characterized by the unit vector  $\mathbf{e}_3$ . We mentioned earlier that at optical frequencies, the fluctuating field components are not observable quantities but quadratic averages of them.

The four instantaneous Stokes parameters at a particular point in space are defined by the sample realizations at time  $t$  of the following combinations of the complex analytic signals  $E_1(t)$  and  $E_2(t)$  associated with the real-valued components of the electric vector in two mutually orthogonal directions perpendicular to  $\mathbf{e}_3$

$$S_0(t) = |E_1(t)|^2 + |E_2(t)|^2 = a_1^2(t) + a_2^2(t) \quad (3.1.112a)$$

$$S_1(t) = |E_1(t)|^2 - |E_2(t)|^2 = a_1^2(t) - a_2^2(t) \quad (3.1.112b)$$

$$\begin{aligned} S_2(t) &= \frac{1}{2}(|E_1(t) + E_2(t)|^2 - |E_1(t) - E_2(t)|^2) = E_1(t)E_2^*(t) + E_1^*(t)E_2(t) \\ &= 2|E_1(t)||E_2(t)| \cos(\theta(t)) = 2a_1(t)a_2(t) \cos(\theta(t)) \end{aligned} \quad (3.1.112c)$$

$$\begin{aligned} S_3(t) &= \frac{1}{2}(|E_1(t) + iE_2(t)|^2 - |E_1(t) - iE_2(t)|^2) = i(-E_1^*(t)E_2(t) + E_1(t)E_2^*(t)) \\ &= 2|E_1(t)||E_2(t)| \sin(\theta(t)) = 2a_1(t)a_2(t) \sin(\theta(t)) \end{aligned} \quad (3.1.112d)$$

The  $S_j$  values form the four-element instantaneous Stokes vector  $\mathbf{S}$ , which contains all the information on the polarization state except the absolute phase of the field, which is seldom of interest.

The usual Stokes parameters are the ensemble averages (or the time average because of our assumption of stationarity and ergodicity) of the instantaneous Stokes parameters. It is important to appreciate at the outset that the very definition of the Stokes parameters appear naturally from consideration of the equation of the polarization ellipse, Eq. (3.1.23). One way to prove this is by

taking the time average of Eq. (3.1.23). It then follows that

$$\frac{\langle E_1^{(r)}(t)^2 \rangle}{a_1^2} + \frac{\langle E_2^{(r)}(t)^2 \rangle}{a_2^2} - 2 \frac{\langle E_1^{(r)}(t)E_2^{(r)}(t) \rangle}{a_1 a_2} \cos(\theta) = (\sin(\theta))^2 \quad (3.1.113)$$

According to the definition of the real orthogonal components of the complex electric field vector, Eq. (3.1.21a,b), we obtain the average values

$$\langle E_1^{(r)}(t)^2 \rangle = \frac{a_1^2}{2} \quad (3.1.114a)$$

and

$$\langle E_2^{(r)}(t)^2 \rangle = \frac{a_2^2}{2} \quad (3.1.114b)$$

If we return to the property of analytic signals as expressed in Eq. (3.1.70a), we have

$$\langle E_j(t)^2 \rangle = a_j^2 \quad (j = 1,2) \quad (3.1.115)$$

Substituting Eqs. (3.1.114a,b) into Eq. (3.1.113), we obtain

$$\langle E_1^{(r)}(t)E_2^{(r)}(t) \rangle = \frac{a_1 a_2 \cos(\theta)}{2} \quad (3.1.116)$$

We again remark, on using Eq. (3.1.70a), that

$$\langle E_1(t)E_2(t) \rangle = a_1 a_2 \cos(\theta) \quad (3.1.117)$$

We left to the reader the task of verifying that

$$(a_1^2 + a_2^2)^2 = (a_1^2 - a_2^2)^2 + (2a_1 a_2 \cos(\theta))^2 + (2a_1 a_2 \sin(\theta))^2 \quad (3.1.118a)$$

or alternatively

$$\sum_{j=1}^3 \langle S_k \rangle^2 = \langle S_0 \rangle^2 \quad (3.1.118b)$$

which is strictly true only for a completely polarized or monochromatic wave.

The time averages of the Stokes parameters are independent of the time instant at which the average is taken by virtue of their being stationary random processes. An important point to bear in mind is that the uniqueness of the set of Stokes parameters is achieved by the introduction of the analytic signal representation. We remind the reader that Stokes vectors for unpolarized,

horizontal, vertical,  $\pm 45^\circ$ , right-hand and left-hand circular states of polarization are listed in Table 3.1.1.

The following fact should also be noted. Stokes parameters have the dimensionality of intensity (or irradiance), which is an International System of Units (SI)-based quantity whose unit is watts per steradian. The prevailing usage is to omit a constant dimensional factor  $\epsilon_0 c/2 = \frac{1}{2}(\epsilon_0/\mu_0)^{1/2}$ , the free-space impedance, common to all four parameters. Consequently, Stokes parameters are dimensionless. Having issued this warning, we shall follow this general practice.

Several conclusions are already emerging. We note first that the orthogonal polarization state

$$\mathbf{E}^* = \begin{bmatrix} E_1^* \\ E_2^* \end{bmatrix}$$

has associated with it the instantaneous Stokes vector

$$\begin{bmatrix} S_0 \\ -S_1 \\ -S_2 \\ -S_3 \end{bmatrix}$$

where the state

$$\mathbf{E} = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix}$$

has associated with it

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix}$$

It is, moreover, easy to see that the first two Stokes parameters involve only the random amplitudes whereas the last two also involve the random phase difference. Equally clear is the fact that the corresponding Stokes vector  $\mathbf{S}_c$  for the circular polarization representation is related to  $\mathbf{S}$  by

$$\mathbf{S}_c = \begin{bmatrix} \langle |E_l|^2 \rangle + \langle |E_r|^2 \rangle \\ \langle |E_l|^2 \rangle - \langle |E_r|^2 \rangle \\ \langle E_l E_r^* \rangle + \langle E_l^* E_r \rangle \\ -i \langle E_l^* E_r \rangle + i \langle E_l E_r^* \rangle \end{bmatrix} = \mathbf{V}_c \mathbf{S} \quad (3.1.119)$$

where

$$\mathbf{V}_c = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} \quad (3.1.120)$$

There are certain important consequences of the definition of the Stokes parameters. While  $S_0$  defines the total instantaneous intensity of the field and is therefore basis-independent, the interpretation of the other Stokes parameters can be made in terms of intensity differences between orthogonal pure states by choosing different basis representations: the excess of parallel to perpendicularly polarized light ( $S_1$ ), the excess of  $+45^\circ$  linearly polarized to  $-45^\circ$  linearly polarized light ( $S_2$ ), and the excess of right circularly to left circularly polarized light ( $S_3$ ) [16].

Thus far the analysis has been confined to a temporal representation of Stokes parameters. A spectral representation of the Stokes parameters is also encountered in the theory of partial polarization [17].

### 3.1.6.2. Coherency (Density) Matrix Formalism

We now devote this section to a discussion of the general formulation of the coherency matrix formalism. The coherency matrix is the covariance matrix of the components  $E_1$  and  $E_2$  of the complex electric field vector at a point specified by position vector  $\mathbf{R}$  in the Euclidian space with basis  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  [18–20]. If  $\mathbf{E}$  is the Jones vector, Eq. (3.1.19), of the wave, the covariance matrix can be constructed from  $\mathbf{E}$ , by writing

$$\Phi_2(0) = \langle \mathbf{E} \otimes \mathbf{E}^+ \rangle \quad (3.1.121)$$

where  $+$   $\equiv (*)^T$  stands for a Hermitian conjugate and the subscript labels the dimension of the matrix. This defines our notation, and allows for comparison with Eq. (3.1.75). (The notation  $\otimes$  stands for the direct product of matrices.) We see that  $\Phi_2(0)$  is a  $2 \times 2$  complex Hermitian nonnegative definite matrix. In subsequent formulas, we drop explicit reference to the argument of  $\Phi_2$ , except when this would lead to ambiguity.

The Stokes parameters at a particular point in space are the coefficients in the expansion of  $\Phi_2$  at that point, in terms of Pauli matrices

$$\Phi_2 = \frac{1}{2} \sum_{j=0}^3 \langle S_j \rangle \sigma_j \quad (3.1.122)$$

where the  $\sigma_j$  values are the usual Hermitian Pauli matrices

$$\sigma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \sigma_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_3 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad (3.1.123a)$$

Observe that we are not restricted to Pauli matrices for an expression of the density matrix. We could clearly use another complete set of four basis matrices for our decomposition

$$\gamma_0 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \gamma_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \gamma_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \gamma_3 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad (3.1.123b)$$

However, the Pauli matrices are the natural ones to use because, as we will show later, they underly the geometry of the Poincaré sphere and provide a fundamental connection between the Mueller and density matrices. It would seem to be useful at this point to keep in mind some well-known identities connecting the Pauli matrices (refer back to footnote 3)

$$\sigma_j \sigma_0 = \sigma_0 \sigma_j = \sigma_j \quad (j \in \langle 0, 3 \rangle) \quad (3.1.124a)$$

$$\sigma_j^2 = \sigma_0 \quad (j \in \langle 0, 3 \rangle) \quad (3.1.124b)$$

the anticommutation relation

$$\sigma_j \sigma_k + \sigma_k \sigma_j = 2\delta_{jk} \sigma_0 \quad (j, k \in \langle 1, 3 \rangle) \quad (3.1.124c)$$

and the commutation relation<sup>6</sup>

$$\sigma_j \sigma_k - \sigma_k \sigma_j = 2i\sigma_l \quad (j, k, l \in \langle 1, 3 \rangle \text{ and cyclic}) \quad (3.1.124d)$$

These matrices are complete and orthogonal, but not normalized since we have

$$\text{tr}(\sigma_j \sigma_k) = 2\delta_{jk} \quad (j, k \in \langle 0, 3 \rangle) \quad (3.1.125)$$

where tr stands for the trace of the matrix that follows.

If Eq. (3.1.122) is both sides multiplied by the matrix  $\sigma_k$ , the trace can be expressed uniquely in the form

$$\text{tr}(\Phi_2 \sigma_k) \equiv \langle S_k \rangle = \langle \sigma_k \rangle \langle S_0 \rangle \quad (k \in \langle 0, 3 \rangle) \quad (3.1.126)$$

Thus the Stokes parameters may be identified with the expectation values of the Pauli matrices  $\sigma_k$  in the polarization state represented by the coherency matrix  $\Phi_2$ .

An alternative to the foregoing analysis is as follows. In many cases the algebraic manipulations are simplified and made more transparent by the introduction of a normalized version of the coherency matrix, rather than the coherency matrix itself:

$$\mathbf{D}_2 \equiv \frac{\Phi_2}{\text{tr}(\Phi_2)} = \frac{1}{2} \sum_{j=0}^3 \langle \sigma_j \rangle \sigma_j \quad (3.1.127)$$

<sup>6</sup> $\sigma_j \sigma_k = \delta_{jk} + i \sum_l \epsilon_{jkl} \sigma_l$ , where  $\delta_{jk}$  is the Kronecker delta symbol and  $\epsilon_{jkl}$  is the Levi-Cevita symbol.

The state of polarization is entirely specified by its density matrix  $\mathbf{D}_2$ . The density matrices of some typical polarization states are given in Table 3.1.1. We note for future reference the important properties of the density matrix: Hermiticity ( $\mathbf{D}_2^+ = \mathbf{D}_2$ ), normalization [ $\text{tr}(\mathbf{D}_2) = 1$ ], and positivity. We do this both for notational simplicity and because in this form observable properties of the field at any point consist of the expectation value of any operator  $\mathbf{O}$  pertaining to the system

$$\langle \mathbf{O} \rangle = \text{tr}(\mathbf{OD}_2) \quad (3.1.128)$$

Therefore, formal manipulations become much simpler with the dimensionless density matrix description of states. Moreover, the density matrix treatment of partial polarization is mathematically equivalent to the description of a two-level system. These remarks offer only the barest indication of why the density matrix formalism has extensive practical applications. The reader who wishes to acquire even an elementary working knowledge of the density matrix theory must consult Ref. 21. We perceive here a first hint of the close association that exists between partial polarization of plane waves and the use of  $SU(2)$ , the group of unitary two-dimensional matrices, an association that we shall explore further in some detail in Section 3.2. In the rest of this volume we shall repeatedly return to this connection as a sort of guiding theme for the description of the partial polarization of a radiation field in the form of plane waves.

Implicit in Eq. (3.1.121) is a choice of a right-hand Cartesian coordinate system of axes. The Jones vector, describing the plane uniform wave propagating along the direction characterized by the unit vector  $\mathbf{e}_3$ , in Eq. (3.1.121) is defined with respect to the linear basis  $(\mathbf{e}_1, \mathbf{e}_2)$  corresponding to orthogonal linear polarizations. However, sometimes it is more convenient to resolve the field into more general orthogonal states of elliptic polarization  $(\mathbf{e}_a, \mathbf{e}_b)$  satisfying the orthonormality condition, i.e.  $\mathbf{e}_i^* \cdot \mathbf{e}_j = \delta_{ij}$ , where  $(i, j) = (a, b)$  and the transversality condition,  $\mathbf{e}_3 \cdot \mathbf{e}_i = 0$ , where  $i = (a, b)$ . One can ask how the expression of the coherency matrix would change if we were to switch to an arbitrary basis  $(\mathbf{e}_a, \mathbf{e}_b)$ . We proceed to exploit this freedom as follows. With respect to this orthonormal basis, the coherency matrix may be expressed as

$$\Phi_2^{(ab)} = \begin{bmatrix} \langle |E_a|^2 \rangle & \langle E_a E_b^* \rangle \\ \langle E_a^* E_b \rangle & \langle |E_b|^2 \rangle \end{bmatrix} \quad (3.1.129)$$

Denoting the transformation from the orthonormal  $(\mathbf{e}_1, \mathbf{e}_2)$  basis to the orthonormal  $(\mathbf{e}_a, \mathbf{e}_b)$  basis by  $\mathbf{U}$ , we have

$$\Phi_2^{(ab)} = \mathbf{U} \Phi_2 \mathbf{U}^{-1} \quad (3.1.130)$$

For instance, we could have chosen the helicity basis ( $\mathbf{e}_l$ ,  $\mathbf{e}_r$ ) to define  $\Phi_2^{(lr)}$  as

$$\Phi_2^{(lr)} = \begin{bmatrix} \langle |E_l|^2 \rangle & \langle E_l E_r^* \rangle \\ \langle E_l^* E_r \rangle & \langle |E_r|^2 \rangle \end{bmatrix} = \mathbf{U}_c \Phi_2 \mathbf{U}_c^{-1} \quad (3.1.131)$$

where the constant matrix  $\mathbf{U}_c$  is defined by Eq. (3.1.20a). It is simple to prove that the trace and the determinant of the coherency matrix are scalar invariants:  $\text{tr}(\Phi_2^{(ab)}) = \text{tr}(\Phi_2)$  and  $\det(\Phi_2^{(ab)}) = \det(\Phi_2)$ .

It is also important to appreciate that the Stokes parameters represent the minimum set of data that are required to completely specify the coherency matrix (quorum of observables). Then, from Eqs. (3.1.122), we arrive at the relation

$$\Phi_2 = \begin{bmatrix} \langle |E_1|^2 \rangle \langle E_1 E_2^* \rangle \\ \langle E_1^* E_2 \rangle \langle |E_2|^2 \rangle \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \langle S_0 \rangle + \langle S_1 \rangle \langle S_2 \rangle - i\langle S_3 \rangle \\ \langle S_2 \rangle + i\langle S_3 \rangle \langle S_0 \rangle - \langle S_1 \rangle \end{bmatrix} \quad (3.1.132)$$

It follows from Eqs. (3.1.127) that the density matrix can be written in the form

$$\mathbf{D}_2 = \frac{1}{2} \begin{bmatrix} 1 + \langle \sigma_1 \rangle & \langle \sigma_2 \rangle - i\langle \sigma_3 \rangle \\ \langle \sigma_2 \rangle + i\langle \sigma_3 \rangle & 1 - \langle \sigma_1 \rangle \end{bmatrix} \quad (3.1.133)$$

By the Cauchy–Schwarz inequality, it follows that the determinant of  $\mathbf{D}_2$  is nonnegative

$$\det(\mathbf{D}_2) \geq 0 \quad (3.1.134)$$

which in the Stokes parameter notation reads  $\sum_{j=1}^3 \langle S_k \rangle^2 \leq \langle S_0 \rangle^2$ . The equality takes place only for totally polarized light. Thus we see that there are three independent Stokes parameters for a pure state of polarization. Values of  $\sum_{j=1}^3 \langle S_k \rangle^2 < \langle S_0 \rangle^2$  identify partially polarized light, that is, nonpure, or mixed states. Remarkably, formula (3.1.133) can be compared with the density matrix of a spin- $\frac{1}{2}$  particle [21].

The trace of the coherency matrix represents the average intensity of the optical field. From the diagonal elements, we can get the intensities while its off-diagonal elements express the correlation between the 1 and 2 components of the Jones vector  $\mathbf{E}$ .

Closely associated with the cross-correlation tensor  $\Phi_2$  is the complex degree of temporal coherence of the light. We have seen that this quantity is defined as the normalized cross-correlation of the electric vibration in the 1 and 2 directions [i.e., Eq. (3.1.79)]:

$$\frac{\Phi_{12}(0)}{(\Phi_{11}(0)\Phi_{22}(0))^{1/2}} \equiv \gamma_{12} = |\gamma_{12}| \exp(i\vartheta_{12}) \quad (3.1.135)$$

Its absolute value measures the degree of correlation of the vibration. By the Cauchy–Schwarz inequality, we have  $0 \leqslant (|\gamma_{12}| \leqslant 1)$ . The lower bound is characteristic of complete (temporal) incoherence and the upper bound by coherence. It is worth observing that the degree of coherence between the electric vibrations in any two mutually orthogonal directions of propagation of the wave depends on the choice of the two orthogonal directions [1].

Up to now we were concerned primarily with the coherency matrix described by space–time correlation functions. However, for many purposes, it is more convenient to characterize them by space–frequency correlation functions. Not to be confused with the usual coherency matrix  $\Phi_2$ , the spectral coherency matrix  $\hat{\Phi}_2$  may be expressed in terms of the spectral Stokes parameters. Assuming that the spectrum is sufficiently narrow to justify the quasimonochromatic approximation, the spectral coherency matrix, at some point specified by the position vector  $\mathbf{R}$ , is related to the cross-spectral tensor by the equation

$$\hat{\Phi}_2 \equiv \lim_{T \rightarrow \infty} \frac{1}{2T} \langle \mathbf{E}(\mathbf{R}, v; T) \times \mathbf{E}^+(\mathbf{R}, v; T) \rangle \quad (3.1.136a)$$

With this expression, we can also define a spectral density matrix as

$$\hat{\mathbf{D}}_2 \equiv \frac{\hat{\Phi}_2}{\text{tr}(\hat{\Phi}_2)} \quad (3.1.136b)$$

The equivalent of Eq. (3.1.136b), but in a more complicated derivation, is given in Ref. 17.

### 3.1.6.3. Spectral Decomposition Theorem

The nonnegative semidefiniteness of coherency matrices  $\Phi_2$  means that their determinants and traces are nonnegative. The purpose of this subsection, then, is to investigate the physical significance of these two parameters. The reader is encouraged to refer to Appendix B for full details. We also refer the interested reader to Refs. 22.

We can derive the physical result by appealing to the spectral decomposition theorem. The simpler version of this theorem is given in Appendix A. For now, we simply want to use Eq. (A.1) and write

$$\Phi_2 = a_0 \sigma_0 + a_1 \mathbf{A}_1 \quad (3.1.137)$$

where  $\mathbf{A}_1$  is an idempotent matrix and  $a_0$  and  $a_0 + a_1$  are the two nonnegative eigenvalues of  $\Phi_2$ . Computation of these eigenvalues  $(1 \pm P)/2\langle S_0 \rangle$  is straightforward and leads to

$$\Phi_2 = \left( \frac{1 - P}{2} \right) \langle S_0 \rangle \sigma_0 + P \langle S_0 \rangle \mathbf{A}_1 \quad (3.1.138)$$

The two terms in this equation have clear physical meanings. This can be appreciated by considering the two limiting cases of unpolarized light ( $P = 0$ , subscript  $u$ ) and totally polarized light ( $P = 1$ , subscript  $p$ ). For unpolarized light, the coherency matrix  $\Phi_{2u}$  is diagonal and proportional to the unit matrix  $\sigma_0$  [i.e.,  $\Phi_{2u} = \langle S_0 \rangle_u (\sigma_0/2)$ ] in agreement with the general result that a density matrix  $\mathbf{D}$  proportional to the unit matrix corresponds to complete ignorance of the state of the system. For totally polarized light, we have  $\Phi_{2p} = \langle S_0 \rangle_p \mathbf{A}_1$ . Thus, a mixed (partially polarized) state can be decomposed as an incoherent mixture of a pure (totally polarized) state and the completely unpolarized state.

It follows immediately from Eq. (3.1.135) that, for completely unpolarized light

$$|\gamma_{12}| = 0 \quad (3.1.139)$$

This means that the electric vibrations in the 1 and 2 directions are mutually incoherent. The corresponding result for pure states is

$$\gamma_{12} = \exp(i\theta) \quad (3.1.140)$$

The absolute value of the complex degree of temporal coherence is unity; its phase is equal to the difference between the phases of the two components. A strictly coherent wavefield is also completely polarized. Another important distinction arises between pure and mixed states. The pure states  $\Phi_{2p} = \langle S_0 \rangle_p \mathbf{A}_{1p}$  have the properties of projectors:

$$(\Phi_{2p})^x = \Phi_{2p} \text{tr}(\Phi_{2p})^{x-1} \quad x \text{ integer}; \quad \det(\Phi_{2p}) = 0 \quad (3.1.141a)$$

or equivalently

$$(\mathbf{D}_{2p})^x = \mathbf{D}_{2p}; \quad \det(\mathbf{D}_{2p}) = 0 \quad (3.1.141b)$$

It has to be noted that the projection property is dropped for a mixed state.

It is important to appreciate that the determinant of  $\Phi_2$  can be rewritten as the following quadratic form

$$\det(\Phi_2) = \frac{1}{4} \langle \mathbf{S} \rangle^+ \mathbf{G} \langle \mathbf{S} \rangle \quad (3.1.142)$$

where  $\mathbf{G}$  is the diagonal Minkowski metric matrix:

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (3.1.143)$$

This expression will be useful for subsequent applications.

From Eq. (3.1.138) we recognize that

$$\frac{\Phi_2}{\langle S_0 \rangle} = (1 - P) \frac{\Phi_{2u}}{\langle S_0 \rangle_u} + P \frac{\Phi_{2p}}{\langle S_0 \rangle_p} \quad (3.1.144)$$

This two-level decomposition is the basic equation of the coherency matrix formalism. In Section 3.2, we shall see that it is at the basis of the fundamental representation of  $SU(2)$  in polarization optics. Two additional remarks are in order. First,  $\mathbf{A}_1$  is given by

$$P\langle S_0 \rangle \mathbf{A}_1 = \frac{1}{2} \begin{bmatrix} P\langle S_0 \rangle + \langle S_1 \rangle & \langle S_2 \rangle - i\langle S_3 \rangle \\ \langle S_2 \rangle + i\langle S_3 \rangle & P\langle S_0 \rangle - \langle S_1 \rangle \end{bmatrix} \quad (3.1.145)$$

where we define the degree of polarization of the light as

$$P \equiv \frac{\text{tr}(P\langle S_0 \rangle \mathbf{A}_1)}{\text{tr}(\Phi_2)} = \frac{\text{tr}(\Phi_{2p})}{\text{tr}(\Phi_2)} = \frac{\langle S_0 \rangle_p}{\langle S_0 \rangle} \quad (3.1.146)$$

It follows that the degree of polarization is defined as the ratio of the average intensity of the completely polarized part of the radiation to the average total intensity. From Eq. (3.1.146) we recognize that  $P$  can be also written in terms of the two rotational invariants of  $\Phi_2$ :

$$P = \left( 1 - 4 \frac{\det(\Phi_2)}{(\text{tr}(\Phi_2))^2} \right)^{1/2} = (1 - 4 \det(\mathbf{D}_2))^{1/2} \quad (3.1.147)$$

Numerical values of the elements of  $\Phi_2$  depend on the choice of the axes of references (for instance, if the 1- and 2-axes are rotated about the direction of propagation of the wave, the coherency matrix will change) but  $\text{tr}(\Phi_2)$  and  $\det(\Phi_2)$  will not change under a rotation of these axes about the direction of propagation of the wave since they are independent scalar invariants of  $\Phi_2$  with respect to unitary transformations. Consequently,  $P$  is independent of the orientation of the coordinate system.

Furthermore  $P$  is bounded and we note that a definition of the degree of polarization, originally suggested by Wolf, introduces the degree of polarization as the maximum modulus of  $|\gamma_{12}|$  [18]. This can be proved as follows.<sup>7</sup>

<sup>7</sup>Wolf has shown that there always exists a pair of mutually orthogonal directions for which the diagonal intensities are equal [18]. Denoting by  $\Theta$  the angle referring this particular pair of directions with the original choice of axes (1–2), the following condition was obtained:

$$\tan(\Theta) = \frac{\Phi_{22}(0) - \Phi_{11}(0)}{2\text{Re}(\Phi_{12}(0))} = -\frac{\langle S_1 \rangle}{\langle S_2 \rangle} = -(\tan(2\psi))^{-1}$$

For this particular pair of directions, the degree of coherence  $|\gamma_{12}|$  of the electric vibrations takes its maximum value equal to  $P$ .

From Eqs. (3.1.135) and (3.1.147), we see that

$$(1 - P^2) = (1 - \gamma_{12}) \left( \frac{\Phi_{11}(0)\Phi_{22}(0)}{(\Phi_{11}(0) + \Phi_{22}(0))/2} \right) \quad (3.1.148)$$

Since the geometric mean of any two positive numbers cannot exceed their arithmetic mean, we arrive at the relation

$$1 \geq P \geq |\gamma_{12}| \geq 0 \quad (3.1.149)$$

which is what we wanted to prove.

For completeness, we indicate some other important expressions that will be employed in the subsequent analysis. We may express the degree of linear polarization of the light by writing

$$P_L = \frac{(\langle S_1 \rangle^2 + \langle S_2 \rangle^2)^{1/2}}{\langle S_0 \rangle} \quad (3.1.150)$$

An alternative characterization of the Stokes parameters is obtained through the use of angles  $\psi$  and  $\chi$ , or  $\alpha$  and  $\theta$ , which coordinate the polarized component by

$$\langle S_1 \rangle = \langle S_0 \rangle P \cos(2\chi) \cos(2\psi) = P \cos(2\alpha) \quad (3.1.151)$$

$$\langle S_2 \rangle = \langle S_0 \rangle P \cos(2\chi) \sin(2\psi) \equiv P \sin(2\alpha) \cos(\theta) \quad (3.1.152)$$

$$\langle S_3 \rangle = \langle S_0 \rangle P \sin(2\chi) = \pm P \sin(2\alpha) \sin(\theta) \quad (3.1.153)$$

The relations (3.1.153a–c) can be inverted, leaving us with

$$P = \frac{(\sum_{j=1}^3 \langle S_j \rangle^2)^{1/2}}{\langle S_0 \rangle} \quad (3.1.154)$$

$$\tan(2\chi) = \frac{\langle S_3 \rangle}{(\langle S_1 \rangle^2 + \langle S_2 \rangle^2)^{1/2}} \quad (3.1.155a)$$

$$\sin(2\chi) = \frac{\langle S_3 \rangle}{\langle S_0 \rangle} = \langle \sigma_3 \rangle \quad (3.1.155b)$$

and

$$\tan(2\psi) = \frac{\langle S_2 \rangle}{\langle S_1 \rangle} = \frac{\langle \sigma_2 \rangle}{\langle \sigma_1 \rangle} \quad (3.1.155c)$$

A geometric interpretation of these relations is considered in some detail in Section 3.2.

### 3.1.6.4. Principle of Optical Equivalence

We now formulate an important theorem, due to Stokes, involving the effect of superposition of pure states. The principle of optical equivalence states that it is impossible to distinguish between various incoherent superpositions of wavefields, having the same frequency, that may together form a beam with identical Stokes parameters. The line of reasoning is as follows.

Let  $E_{1k}, E_{2k}$ ,  $k \in \langle 1, N \rangle$  denote the analytic signals that represent the components of the electric vibrations of the  $k$ th wave in a plane perpendicular to  $\mathbf{e}_3$ . We assume now that the  $N$  independent beams propagating in the same direction are superposed in space. The components of the resulting wave are

$$E_1 = \sum_{k=1}^N E_{1k} \quad (3.1.156)$$

$$E_2 = \sum_{k=1}^N E_{2k} \quad (3.1.156)$$

If we substitute from Eq. (3.1.156) into Eq. (3.1.121), we obtain the elements of the coherency matrix, because averaging is a linear operation:

$$\Phi_{2ij} = \langle E_i E_j^* \rangle = \sum_{k=1}^N \sum_{l=1}^N \langle E_{ik} E_{jl}^* \rangle = \sum_{k=1}^N \langle E_{ik} E_{jk}^* \rangle + \sum_k \sum_{l \neq k} \langle E_{ik} E_{jl}^* \rangle \quad (3.1.157)$$

Since the waves are assumed to be independent of each other with respect to their amplitudes and phases, each cross-term of the second term of the right-hand side of Eq. (3.1.157) is zero and it follows that

$$\Phi_2 = \sum_{k=1}^N \Phi_{2k} \quad (3.1.158)$$

Thus, if  $N$  mutually independent quasimonochromatic beams that are propagated in the same direction are superposed incoherently, the coherency matrix of the resulting wave is equal to the sum of the coherency matrices of the individual waves. Conversely, any partially polarized beam may be regarded as made up of an incoherent superposition of any number of independent beams that may be chosen arbitrarily except for a few simple cases.

Let  $\mathbf{D}_{2k}$  represent the density matrix for the  $k$ th light beam. It follows at once from Eq. (3.1.158) that the density matrix  $\mathbf{D}_2$  for the global mixture may be expressed as

$$\mathbf{D}_2 = \sum_{k=1}^N \lambda_k \mathbf{D}_{2k} \quad (3.1.159)$$

where the  $\lambda_k$  values with  $k \in \langle 1, N \rangle$  are the intensity weights [ $\lambda_k = \text{tr}(\Phi_{2k})/\text{tr}(\Phi_2)$  and  $\sum_{k=1}^N \lambda_k = 1$ ] of the different components of the mixture. It should be noted that when  $N$  quasimonochromatic light beams (of the same mean frequency) propagating along the same direction are incoherently combined, the degree of polarization of the resultant beam can never be greater than that of the constituent beams. This theorem follows directly from Eq. (3.1.159); see Appendix C for the details of the proof.

Similarly, one can write the resulting Stokes parameters as

$$\langle S_j \rangle = \sum_{k=0}^N \langle S_j \rangle_k \quad (3.1.160)$$

Thus, an important property of the Stokes parameters is their additivity for an incoherent superposition of waves.

The following decomposition is also useful. Any arbitrary partially polarized state can always be decomposed into an elliptically polarized part and an unpolarized part:

$$\mathbf{S} \equiv \begin{bmatrix} \langle S_0 \rangle - \left( \sum_{j=1}^3 \langle S_j \rangle^2 \right)^{1/2} \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \left( \sum_{j=1}^3 \langle S_j \rangle^2 \right)^{1/2} \\ \langle S_1 \rangle \\ \langle S_2 \rangle \\ \langle S_3 \rangle \end{bmatrix} \quad (3.1.161)$$

The orientation of the principal axes and the ellipticity of the polarized component are given by Eqs. (3.1.155a–c). This expression is important because it parallels Eq. (3.1.144), albeit in terms of Stokes parameters. A discussion of the geometric of these analytic formulas is postponed until Section 3.2.

Another remarkable result connected with the two-level description of polarization states is as follows. Any arbitrary state of polarization is an incoherent superposition of two orthogonally elliptically polarized components

$$\mathbf{S} \equiv \frac{1+P}{2P} \begin{bmatrix} P\langle S_0 \rangle \\ \langle S_1 \rangle \\ \langle S_2 \rangle \\ \langle S_3 \rangle \end{bmatrix} + \frac{1-P}{2P} \begin{bmatrix} P\langle S_0 \rangle \\ -\langle S_1 \rangle \\ -\langle S_2 \rangle \\ -\langle S_3 \rangle \end{bmatrix} \quad (3.1.162)$$

where the degree of polarization is given by Eq. (3.1.154).

Unpolarized light can be obtained (non uniquely) by incoherently superposing any two orthogonally pure states of polarization in equal proportions

$$\langle S_0 \rangle \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \frac{\langle S_0 \rangle}{2} \begin{bmatrix} 1 \\ \cos(2\chi) \cos(2\psi) \\ \cos(2\chi) \sin(2\psi) \\ \sin(2\chi) \end{bmatrix} + \frac{\langle S_0 \rangle}{2} \begin{bmatrix} 1 \\ -\cos(2\chi) \cos(2\psi) \\ -\cos(2\chi) \sin(2\psi) \\ -\sin(2\chi) \end{bmatrix} \quad (3.1.163)$$

opposite ellipticities ( $\chi' = -\chi$ ) and polarization ellipses oriented at  $\pi/2$  from each other ( $\psi' = \psi + (\pi/2)$ ). For example, unpolarized light can be decomposed into an incoherent sum of linear horizontal and linear vertical states of polarization:

$$\langle S_0 \rangle \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \frac{\langle S_0 \rangle}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \frac{\langle S_0 \rangle}{2} \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} \quad (3.1.164)$$

Note that when the two pure states of polarization are not mixed in equal proportions, we find that the resulting state is partially polarized.

### 3.1.6.5. Reduction Property

In general, the second-order correlation tensor  $\Phi_2(\mathbf{R}_1, \mathbf{R}_2, \tau)$  has a rather complex structure. In deriving the wave equations for this correlation tensor, we have already seen that spatial and temporal coherence properties of waves are not independent of each other. However, under certain circumstances, the tensor  $\Phi_2(\mathbf{R}_1, \mathbf{R}_2, \tau)$  at two points  $(\mathbf{R}_1, \mathbf{R}_2)$  in an optical field is reducible to the product of a function of  $\mathbf{R}_1, \mathbf{R}_2$  alone and a function of the time difference  $\tau$ . This is called the *reduction property*, a concept first introduced by Mandel [23]. A radiation field which possesses this reduction property will be called a *spectrally pure beam of light*. Light beams that do not have this property are called spectrally impure. Reference is made to Mandel [23] and Mandel and Wolf [1, 24] for details of the derivation of these conditions.

Since partial polarization deals with the correlation of the electric field vector at any point  $\mathbf{R}$ , this subsection is devoted to cover a particularly interesting situation, namely, the conditions under which  $\Phi(\mathbf{R}, \mathbf{R}, \tau)$  may be expressed as a product of a function of  $\tau$  alone and of the coherency matrix  $\Phi_2(0)$ . As we have done previously, we again assume that the optical field is quasimonochromatic and stationary. The expression of  $\Phi_2$  may be written in

matrix form as

$$\Phi_2(\tau) = \begin{bmatrix} \langle |E_1(t)|^2 \rangle \gamma_{11}(\tau) & (\langle |E_1(t)|^2 \rangle \langle |E_2(t)|^2 \rangle)^{1/2} \gamma_{12}(\tau) \\ (\langle |E_1(t)|^2 \rangle \langle |E_2(t)|^2 \rangle)^{1/2} \gamma_{21}(\tau) & \langle |E_2(t)|^2 \rangle \gamma_{22}(\tau) \end{bmatrix} \quad (3.1.165)$$

where we have used the normalization convention, Eq. (3.1.78). Next let us assume that

$$\gamma_{11}(\tau) = \gamma_{22}(\tau) \quad (3.1.166a)$$

and

$$\gamma_{12}(\tau) = \gamma_{12}(0)\gamma_{11}(\tau) \quad (3.1.166b)$$

Under these circumstances  $\Phi_2$  can now be written as

$$\Phi_2(\tau) = \gamma_{11}(\tau)\Phi_2(0) \quad (3.1.167)$$

The reduction formula, Eq. (3.1.167), asserts that  $\Phi_2(\tau)$ , which characterizes temporal coherence at some point, is expressible as the product of two factors: one factor is the normalized correlation factor  $\langle E_1(t)E_1^*(t + \tau) \rangle / \langle |E_1(t)|^2 \rangle$ , and the other is the coherency matrix  $\Phi_2(0)$ . An important feature to emerge from this analysis is simply the observation that, if the factorization formula Eq. (3.1.167) is satisfied, the degree of polarization calculated from Eq. (3.1.147) is independent of the time delay  $\tau$  [i.e.,  $P(\tau) = P(0)$ ]. It is a consequence of Eq. (3.1.78) that  $\text{tr}(\Phi_2(\tau)) \leq \text{tr}(\Phi_2(0))$ . Finally, the reader is reminded that Eq. (3.1.167) is a special case of the more general result given by Mandel and Wolf [1].

### 3.1.6.6. Covariance Matrix of Instantaneous Stokes Parameters

While interest has usually been focused on second-order correlations, fourth-order measurements are both practical and useful; examples are the Hanbury-Brown and Twiss intensity correlations experiments [2, 25]. This section is devoted to the analysis of covariance matrix of the instantaneous Stokes parameters. Borrowing from our experience with the second-order description of the fluctuations  $2 \times 2$  in terms of the coherency matrix  $\Phi_2$ , we now pass to a consideration of a  $4 \times 4$  coherency matrix  $\Phi_4$ . We have chosen to characterize this coherency matrix as  $\Phi_4 = \langle \mathbf{S} \otimes \mathbf{S}^+ \rangle - \langle \mathbf{S} \rangle \otimes \langle \mathbf{S}^+ \rangle$  rather than  $\langle \mathbf{S} \otimes \mathbf{S}^+ \rangle$ . It is because the former yields a coherency matrix proportional to the  $4 \times 4$  unit matrix  $\mathbf{O}_0^{(4)}$  for polarized light as required. Whereas the components of  $\Phi_2$  are complex and involve second-order moments  $\langle E_p^* E_q \rangle$  of the field the reader can see, the components of  $\Phi_4$  are real and related to the fourth-order moments  $\langle E_p^* E_q^* E_r E_s \rangle$  of the field. So far, no assumption has been made about the PDF that governs the fluctuations. It will now be assumed that this PDF is Gaussian. With the help of the moment reduction

formula<sup>8</sup> for a complex Gaussian random process, the fourth-order field correlations can be decomposed into a sum of products of second-order moments. It is instructive to apply again the spectral decomposition theorem (with reference to Appendix B) in making the calculation of  $\Phi_{4g}$ , where the subscript  $g$  stands for Gaussian. A quick calculation along the lines of Eq. (A.1) yields the simple form

$$\Phi_{4g} = \frac{1}{2} \begin{bmatrix} \left(\frac{1+P^2}{2}\right)\langle S_0 \rangle^2 & \langle S_0 \rangle \langle S_1 \rangle & \langle S_0 \rangle \langle S_2 \rangle & \langle S_0 \rangle \langle S_3 \rangle \\ \langle S_0 \rangle \langle S_1 \rangle & \left(\frac{1-P^2}{2}\right)\langle S_0 \rangle^2 + \langle S_1 \rangle^2 & \langle S_1 \rangle \langle S_2 \rangle & \langle S_1 \rangle \langle S_3 \rangle \\ \langle S_0 \rangle \langle S_2 \rangle & \langle S_1 \rangle \langle S_2 \rangle & \left(\frac{1-P^2}{2}\right)\langle S_0 \rangle^2 + \langle S_2 \rangle^2 & \langle S_2 \rangle \langle S_3 \rangle \\ \langle S_0 \rangle \langle S_3 \rangle & \langle S_1 \rangle \langle S_3 \rangle & \langle S_2 \rangle \langle S_3 \rangle & \left(\frac{1-P^2}{2}\right)\langle S_0 \rangle^2 + \langle S_3 \rangle^2 \end{bmatrix} \quad (3.1.168)$$

We see that the Gaussian statistics has for effect to decorrelate the Stokes parameters; consequently, the description of polarization states by  $\Phi_2$  is equivalent to that given by  $\Phi_{4g}$ . Equation (3.1.168) can be recast in a compact form as

$$\Phi_{4g} = \left(\frac{P^2 - 1}{4}\right)\langle S_0 \rangle^2 \mathbf{G} + \frac{1}{2}\langle \mathbf{S} \rangle \otimes \langle \mathbf{S} \rangle^+ \quad (3.1.169)$$

Even though this equation may be considered our final result, a more elegant formula may be obtained

$$\Phi_{4g} = \mathbf{A}(\Phi_{2g} \otimes \Phi_{2g})\mathbf{A}^{-1} \quad (3.1.170)$$

where the constant matrix  $\mathbf{A}$  ( $\mathbf{A}^{-1} = \frac{1}{2}\mathbf{A}^+$ ) is of the form

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \quad (3.1.171)$$

We will utilize this result in our discussion of the Mueller formalism to deal

<sup>8</sup>A well-known result in random process theory is that if  $A$ ,  $B$ ,  $C$ , and  $D$  are real random processes with a Gaussian joint PDF and each with a zero mean, we have  $\langle ABCD \rangle = \langle AB \rangle \langle CD \rangle + \langle AC \rangle \langle BD \rangle + \langle AD \rangle \langle BC \rangle$  [4].

with the effect of a non-image-forming optical system (or a scattering medium) on an incident partially polarized plane wave. A more useful formulation of  $\Phi_{4g}$  follows from the spectral decomposition theorem developed in Appendix B. Consequently the problem shifts to the evaluation of the eigenvalues of  $\Phi_{4g}$ . We find  $[(1 - P^2)/4] \langle S_0 \rangle^2$ , which is doubly degenerate,  $[(1 + P)/2]^2 \langle S_0 \rangle^2$ , and  $[(1 - P)/2]^2 \langle S_0 \rangle^2$ . One can take advantage of Eq. (A.1) to write  $\Phi_{4g}$  as

$$\begin{aligned}\Phi_{4g} = a_0 \mathbf{O}_0^{(4)} + a_1 \mathbf{A}_1 + a_2 \mathbf{A}_2 &= \left( \frac{1 - P^2}{4} \right) \langle S_0 \rangle^2 \mathbf{O}_0^{(4)} \\ &+ \frac{P(P + 1)}{2} \langle S_0 \rangle^2 \mathbf{A}_1 + \frac{P(P - 1)}{2} \langle S_0 \rangle^2 \mathbf{A}_2\end{aligned}\quad (3.1.172)$$

where  $\mathbf{A}_j$  are two idempotent matrices that can be easily evaluated.

From the expression of the rotational invariants of  $\Phi_{4g}$ , we leave to the reader the task of proving that the degree of polarization takes the form

$$P = \left( 1 - 4 \frac{(\det(\Phi_{4g}))^{1/4}}{\text{tr}(\Phi_{4g})} \right)^{1/2} = (1 - 4 \det(\mathbf{D}_{4g})^{1/4})^{1/2} \quad (3.1.173)$$

which greatly resembles Eq. (3.1.75).

An interesting problem would be the extension of the analysis to consider the general (non-Gaussian PDF) case, where one would have to deal with the properties of the fourth-order correlation tensor  $\Phi_4$  without the simplification afforded by Eq. (3.1.168). This problem certainly deserves further work because intensity fluctuations are much slower than amplitude oscillations. This concludes our discussion of the coherency (density) matrix formalism. The results that we discussed in this section will be useful to us in Part 4. Our attention is now turned to a presentation of generalized Stokes parameters and of a generalized density matrix that may be useful to deal with the polarization of nonplane waves.

### 3.1.6.7. Digression: Generalized Stokes Parameters

In this subsection we indicate briefly how the density matrix approach to partial polarization of plane waves of infinite extent discussed in the preceding subsections can be generalized to deal with waves of arbitrary form. We are concerned with a narrowband optical field that can be represented by an ensemble of realizations, which we shall assume to be statistically stationary, at least in the wide sense. The key mathematical idea utilized here is that the available information on the wavefield is the density matrix that describes the second-order temporal correlation between the electric field components at a given point in space. So far no assumption has been made about the statistics governing the light fluctuations; we limit our description of the statistical properties of the underlying wavefields to second order.

By definition,  $\mathbf{D}_N$  is nonnegative definite and Hermitian:  $\mathbf{D}_N$  can be diagonalized by a unitary transformation, and its  $N$  eigenvalues are real and nonnegative. On the basis of this description and by analogy with the case  $N = 2$ , we introduce the normalized Stokes parameters  $\Theta_j^{(N)}$ , which are defined by the scalar coefficients in the expansion of  $\mathbf{D}_N$  in terms of the  $N^2$  Hermitian, trace orthogonal, and linearly independent  $\mathbf{O}_j^{(N)}$  matrices. An important point to appreciate here is that these real parameters form a quorum of observables that completely specify the state of polarization of the optical field. Let us introduce formally the  $N \times N$  polarization density matrix as a linear combination of  $N^2$  independent Hermitian matrices

$$\mathbf{D}_N = \frac{1}{N} \sum_{j=0}^{N^2-1} \Theta_j^{(N)} \mathbf{O}_j^{(N)} \quad (3.1.174)$$

This equation is known as the  $SU(N)$  expansion of the polarization density matrix  $\mathbf{D}_N$  and generalizes Eq. (3.1.127), but with the Pauli matrices now being replaced by the generators of the group  $SU(N)$ . It is convenient to work with a normalized version of the Stokes parameters, say,  $\Theta_j^{(N)}$ , since they take a dimensionless form. The expectation value of a physical observable, characterizing the light at any point, described by the density matrix  $\mathbf{D}_N$  is given by

$$\langle \mathbf{O}_j^{(N)} \rangle \equiv \text{tr}(\mathbf{O}_j^{(N)} \mathbf{D}_N) = \Theta_j^{(N)} \quad (3.1.175)$$

where the angular brackets denote the average taken over the statistical ensemble representing the fluctuating field. Two important points should be stressed. First, we set aside  $\mathbf{O}_0^{(N)}$  to be the  $N \times N$  identity matrix. The second important feature is the trace relations, namely, the normalization condition

$$\text{tr}(\mathbf{O}_j^{(N)}) = N\delta_{j0} \quad (3.1.176a)$$

and the orthogonality condition

$$\text{tr}(\mathbf{O}_j^{(N)} \mathbf{O}_k^{(N)}) = 2\delta_{jk} \quad (3.1.176b)$$

For  $N = 2$ , we have previously seen that the  $\mathbf{O}_j$  terms are the Pauli matrices and  $\Theta_j^{(2)} = \langle \sigma_j \rangle$ . For  $N = 3$ , the  $\mathbf{O}_j$  terms are the Gell–Mann matrices and for  $N = 4$ , the Dirac matrices. At this point it is worth commenting on several properties of  $\mathbf{D}_N$ . Of particular importance in the analysis is the concept of scalar invariants. Scalar functions of  $\mathbf{D}_N$  that are invariant, with respect to all transformations,  $\mathbf{D}_N \rightarrow \mathbf{U} \mathbf{D}_N \mathbf{U}^{-1}$ , where  $\mathbf{U}$  is a unitary matrix are the scalar invariants of  $\mathbf{D}_N$ . Simple examples are the trace and the determinant of  $\mathbf{D}_N$ . The Cayley–Hamilton theorem implies that these basic invariant quantities can be obtained by direct evaluation of the traces of powers of  $\mathbf{D}_N$ . The physical significance of these scalar invariants as measures of the degree of polarization of the optical field is fully discussed in Refs. 27 and 28.

Let us examine the meaning of this parametrization for  $N = 3$ . As already mentioned, for  $N = 3$ , the nine linearly independent  $3 \times 3$  matrices Lie group  $SU(3)$ ,  $\{\mathbf{O}_j^{(3)}\}$ , forms the Gell–Mann set. The nine matrices  $\mathbf{O}_j^{(3)}$  form a complete set of  $3 \times 3$  matrices

$$\mathbf{O}_0^{(3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{O}_1^{(3)} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{O}_2^{(3)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (3.1.177a)$$

$$\mathbf{O}_3^{(3)} = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{O}_4^{(3)} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad \mathbf{O}_5^{(3)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (3.1.177b)$$

$$\mathbf{O}_6^{(3)} = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}, \quad \mathbf{O}_7^{(3)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \quad \mathbf{O}_8^{(3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (3.1.177c)$$

The preceding discussion assumes that the light was in the form of plane waves. Now the spatial structure of the quasimonochromatic wave is assumed to be arbitrary; an example would be the field in the focal volume of a lens illuminated by a plane wave. Contrasting with plane waves that can be described, at a fixed point in space, by two transverse components of the electric field, the case of nonplanar waves requires to consider a longitudinal component  $E_3$ . Consequently, the polarization density matrix is a  $3 \times 3$  matrix that may be written in the form

$$\mathbf{D}_3 = \frac{1}{\sum_{j=1}^3 \langle |E_j|^2 \rangle} \begin{bmatrix} \langle |E_1|^2 \rangle & \langle E_1 E_2^* \rangle & \langle E_1 E_3^* \rangle \\ \langle E_1^* E_2 \rangle & \langle |E_2|^2 \rangle & \langle E_2 E_3^* \rangle \\ \langle E_1^* E_3 \rangle & \langle E_2^* E_3 \rangle & \langle |E_3|^2 \rangle \end{bmatrix} \quad (3.1.178a)$$

and is determined by eight independent measurements. Using the expansion of  $\mathbf{D}_3$  in terms of a linear combination of Gell–Mann matrices, an equivalent expression in terms of the normalized Stokes parameters is the following:

$$\mathbf{D}_3 = \frac{1}{3} \begin{bmatrix} 1 + \frac{\Theta_4^{(3)}}{\sqrt{3}} + \Theta_3^{(3)} & \Theta_1^{(3)} - i\Theta_2^{(3)} & \Theta_5^{(3)} - i\Theta_6^{(3)} \\ \Theta_1^{(3)} + i\Theta_2^{(3)} & 1 + \frac{\Theta_4^{(3)}}{\sqrt{3}} - \Theta_3^{(3)} & \Theta_7^{(3)} - i\Theta_8^{(3)} \\ \Theta_5^{(3)} + i\Theta_6^{(3)} & \Theta_7^{(3)} + i\Theta_8^{(3)} & 1 - \frac{2\Theta_4^{(3)}}{\sqrt{3}} \end{bmatrix} \quad (3.1.178b)$$

The expressions of the  $\Theta_j^{(3)}$  values in terms of the covariances of the electric

vector components are listed in Appendix D. An argument similar to that of Section 3.1.5.2 suggests that there exists now an analogy of the analysis of polarization for a nonplanar wavefield with the statistical description of a spin-1 particle.

For  $N = 2$ , we recall that the degree of polarization is given by Eq. (3.1.147). In this case we can show that the only order parameter is the rotational invariant  $\text{tr}((\mathbf{D}_2)^2) = 1 - 2 \det(\mathbf{D}_2)$ . For  $N = 3$ , two order parameters are required. Consequently two degrees of polarization may be defined in terms of the scalar invariants  $\text{tr}((\mathbf{D}_3)^2)$  and  $\text{tr}((\mathbf{D}_3)^3) = 3 \det(\mathbf{D}_3)$ . As we have done previously, we again consider the extreme cases of unpolarized and completely polarized waves. Let us first consider the case of unpolarized light. Observing that  $\mathbf{D}_3 = \frac{1}{3}\mathbf{O}_0^{(3)}$ , we see that  $\text{tr}((\mathbf{D}_3)^2) = \frac{1}{3}$  and  $\text{tr}((\mathbf{D}_3)^3) = \frac{1}{9}$ . On the other hand, it is worth noting that,  $\det(\mathbf{D}_3) = 0$  is not the necessary and sufficient condition that the wave is in a pure state; one also requires that  $\text{tr}(\mathbf{D}_3^2) = \frac{1}{4}$ . Another interesting fact concerns the situation for which  $E_3 = 0$ . In that case, we find that

$$\mathbf{D}_3 = \begin{bmatrix} \mathbf{D}_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

where  $\mathbf{0}$  is a  $3 \times 3$  diagonal matrix with all entries being equal to zero, and we have  $\text{tr}((\mathbf{D}_3)^3) = 0$  and  $\text{tr}((\mathbf{D}_3)^2) = \text{tr}((\mathbf{D}_2)^2)$ ; that is, we recover the case of plane waves oriented arbitrarily with respect to a specified coordinate system.

Several remarks pertinent to the matters at hand should be mentioned. While pure and mixed states could be easily positioned relative to each other in the convex set of states for  $N = 2$ , no such simple topology exists for  $N = 3$  because the  $N = 3$  set has too many dimensions. Indeed, we find that the convex set of states has three strata with dimensions 4, 7, and 8 leading to a complex hierarchy of hypersurfaces. For  $N = 2$ , the set of pure states is identical with the topology boundary of the domain of positivity within the set of all Hermitian  $N \times N$  matrices of unit trace. For  $N > 2$ , this boundary presents more complex topological properties involving curves edges and rulings of various dimensions. Another difference arises from a simple parameter count. For  $N = 2$ , it was possible to consider an arbitrary (partially polarized) state as an incoherent mixture of a pure state (three independent parameters) and the completely unpolarized state. However, for  $N = 3$ , we need eight independent quantities to specify the density matrix  $\mathbf{D}_3$ , and consequently we cannot generally express an arbitrary state as a mixture of a pure state (five independent parameters) and the completely unpolarized state. Obvious physical applications of the case  $N = 3$  are to problems concerning near-field situations, such as polarization diffraction microscopy.

This concludes our discussion of the basic concepts underlying the theory of partial polarization. We shall find many of the concepts and results of this chapter reappearing in subsequent discussions of partially polarized light

interacting with optical media. Three other aspects will be dealt in subsequent sections; in Section 3.2, we turn our attention to some geometric representations of partially polarized light. In Section 3.3, we describe some features of the formalism appropriate for the statistics of the radiation field. We pay special attention to the case of Gaussian light. Finally, in Section 3.4, we describe the physics underlying the description of the entropy of the radiation field.

## SECTION 3.2

---

# Geometric Representations of Partially Polarized Light

Thus far only the algebraic parametrizations of the states of polarization of an optical field were described and exploited. In this Section we have a twofold purpose. On one hand, we shall attempt to describe different geometric methods of representing the state of polarization of plane waves, such as the stereographic projection and the Poincaré sphere. These methods both achieve the same end; however, whereas the former does so by means of transformations in a complex two-dimensional space, the latter utilizes transformations in three-dimensional space. Like all concepts, we understand them best when they are expressed in geometric terms. On the other hand, we shall investigate the interrelation of the algebraic and geometric approaches. The unification of approach is suggested by the application of  $SU(2)$  labeling to plane-wave polarization.

The body of this section is organized as follows. Section 3.2.1 is devoted to a summary of those aspects of Stokes vector space directly pertinent to our purpose. The Poincaré sphere, which is the usual sphere of unit radius in  $R^3$ , provides an intuitively geometric view for the components of the Stokes vector  $S$ . The advantage of introducing the Stokes vector space is that it allows us to describe the state of polarization for both completely and partially polarized wavefields in geometrical terms. Section 3.2.2 discusses the convexity structure of sets of polarization states. Our guiding theme, here, is that the set of pure states is rich enough to generate, by convex combination, the set of all mixed states. The Poincaré sphere may alternatively be parametrized by the coordinates on the equatorial plane by stereographic projection that establishes a correspondence between the points of a plane and those of  $\Sigma_1^2$ . Section 3.2.3 covers this aspect of the problem in some detail and describes the principle of stereographic projection. Geometric phase forms the subject of Section 3.2.4. The  $SU(2)$  algebra arises naturally when one adopts the two-level description of partially polarized light. Finally in Section 3.2.5, the  $SU(2)$  formalism provides us with a geometric overview of the equation of motion for the Poincaré vector (the Bloch vector). Throughout this chapter, the convention used is to define the polarization state in the linear basis, defined by  $e_1$  and  $e_2$  unit vectors. Such a choice is arbitrary and does not have any physical significance, but once it is made, we can identify the physical effects by recalling

that directions 1 and 2 in the Stokes vector space correspond to linear polarizations and direction 3, to circular polarization.

### 3.2.1. THE STOKES VECTOR SPACE

The use of geometric ideas in polarization optics dates back to Poincaré. When Poincaré published his famous 1892 treatise, *Théorie Mathématique de la Lumière*, he gave a remarkable example of the relationship between geometry and measurable physical quantities such as Stokes parameters [29]. In fact, the whole book is permeated by geometric ideas. In the Poincaré sphere representation, any pure state of polarization can be represented by a point

$$M \begin{bmatrix} \langle \sigma_1 \rangle \\ \langle \sigma_2 \rangle \\ \langle \sigma_3 \rangle \end{bmatrix}$$

on the surface of a sphere of unit radius. As this surface is two-dimensional, each point may be coordinated by spherical angles  $2\psi$  and  $2\chi$  for its specification. The latitude and longitude of each point are defined by the ellipticity and azimuth angles of the polarization ellipse. The correspondence between the point  $M$  and the various states of polarization is most easily seen using the Poincaré sphere, a surface in Stokes vector space (see Fig. 3.2.1).

The unit Poincaré vector  $u(\chi, \psi)$  allows a clear geometric interpretation. The points on the equator  $0 \leq 2\psi \leq 2\pi$ ,  $2\chi = 0$  represent linearly polarized light. Vectors

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$$

are the horizontal and vertical linear polarizations. The vector

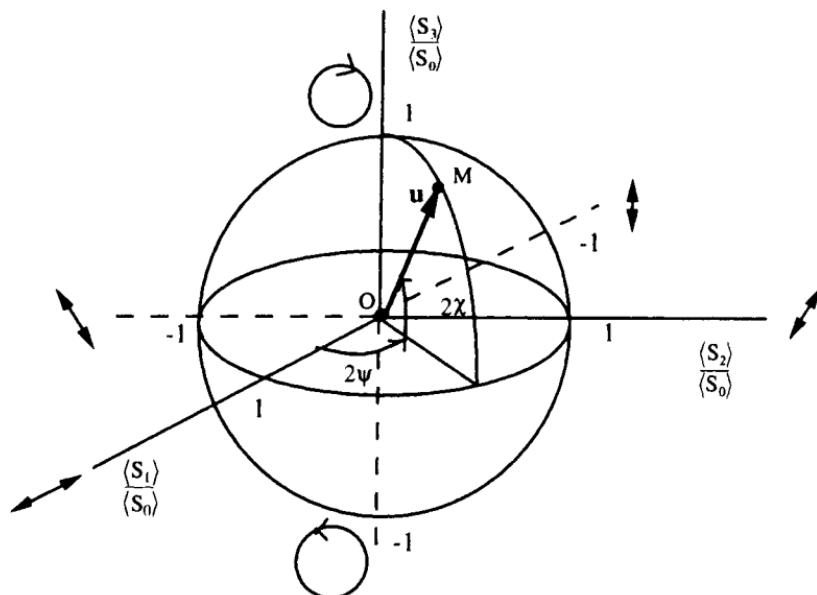
$$\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

points to the north pole

$$N \left( 2\chi = \frac{\pi}{2} \right)$$

of  $\Sigma_2^1$  and represents right-handed circular polarization, while

$$\begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$



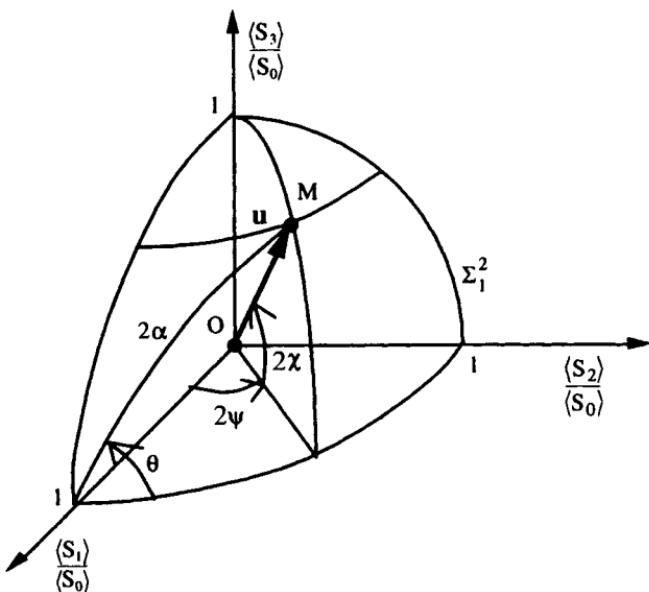
**Figure 3.2.1.** The Poincaré sphere  $\Sigma^2$  is the unit sphere surrounding the origin of the Cartesian coordinate, orthonormal basis  $(\mathbf{e}_1, \mathbf{e}_2)$ . The normalized Stokes parameters  $(\langle \sigma_1 \rangle, \langle \sigma_2 \rangle, \langle \sigma_3 \rangle)$  constitute the components of the Poincaré vector  $\mathbf{u}$  that represents the state of polarization of an arbitrary pure state of polarization ( $|\mathbf{u}| = 1$ ). The longitude  $2\psi$  and latitude  $2\chi$  of point  $M$  are respectively related to the azimuth and the ellipticity angles of the polarization ellipse of the wave. Each point on  $\Sigma^2$  corresponds to a unique state of polarization. The north pole  $N = [0, 0, 1]^T$  represents right circularly polarized light. The south pole  $S = [0, 0, -1]^T$  represents left circularly polarized light. Points on the equator ( $2\chi = 0$ ) represent linearly polarized light. Elliptical polarization states lie between the poles and equator. The positive directions of the angle  $2\psi$  and  $2\chi$  are defined according the adopted sign convention.

points to the south pole  $S$  ( $2\chi = -\pi/2$ ) and represents left-handed circular polarization.

In Fig. 3.2.2, the Poincaré vector  $\mathbf{u}$  described by point  $M$  on the Poincaré sphere  $\Sigma^1$  can be shown in terms of angles  $\alpha$  and  $\theta$ :  $2\alpha$  is the angle subtended by the great circle drawn from point  $M$  on the equator and  $\theta$  is the angle between the great circle and the equator.

To summarize, we have succeeded in representing the parameters of the polarization ellipse on the Poincaré sphere by spherical trigonometry. Figure 3.2.2 helps us visualize the relations that exist between the various angles that characterize the polarization ellipse, given previously by Eqs. (3.1.151)–(3.1.153).

Points on the poles represent the two opposite circularly polarized states. All other points lying between the poles and equator describe elliptically



**Figure 3.2.2.** Schematic illustration of the relation of angles  $\chi$ ,  $\psi$ ,  $\alpha$ , and  $\theta$ .

polarized states; their positions are determined by the ellipticity of the state of polarization and the orientation of the major axis of the ellipse. Right elliptical polarization states are located in the northern hemisphere, whereas points on the southern hemisphere represent left-elliptical polarization states. Antipodal points on the sphere represent orthogonal polarization states, that is, opposite ellipticities and with azimuth differing by  $\pi/2$ , for example, one polarization right-circular and the other left-circular. Points that lie on the same parallel ( $2\chi = \text{constant}$ ) represent all possible forms of elliptical polarization with the same ellipticity but with different azimuths. Points that lie on the same meridian ( $2\psi = \text{constant}$ ) represent all possible forms of elliptical polarization with the same azimuth but with different ellipticities.

The Poincaré method is applicable to any state of polarization. Specifically, it is easy to show that the Stokes vector can be expressed as

$$\mathbf{S} \equiv \langle S_0 \rangle \begin{bmatrix} 1 \\ \mathbf{P}u(\chi, \psi) \end{bmatrix}$$

Such a parametrization is helpful because it applies for both pure and mixed states of polarization. Hence, the points inside the unit ball  $\Sigma_3^1$  represent the polarization states of partially polarized fields (mixed states).

An important theorem is the following. Any arbitrary Poincaré vector can always be written as a sum of three pure states. The proof goes as follows.

Partially polarized light can be represented as a sum of a purely polarized and unpolarized components with weights  $P$  and  $1-P$  assigned to each. The unpolarized part, in turn, can be viewed as an incoherent sum of orthogonally pure states of equal intensity:

$$\begin{aligned}\langle S_0 \rangle \begin{bmatrix} 1 \\ \mathbf{P}\mathbf{u}(\chi, \psi) \end{bmatrix} &= P\langle S_0 \rangle \begin{bmatrix} 1 \\ \mathbf{u}(\chi, \psi) \end{bmatrix} + (1 - P)\langle S_0 \rangle \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ &= P\langle S_0 \rangle \begin{bmatrix} 1 \\ \mathbf{u}(\chi, \psi) \end{bmatrix} + (1 - P)\langle S_0 \rangle \left( \begin{bmatrix} 1 \\ \mathbf{v}(\chi', \psi') \end{bmatrix} + \begin{bmatrix} 1 \\ -\mathbf{v}(\chi', \psi') \end{bmatrix} \right)\end{aligned}\quad (3.2.1)$$

Note that the choice of the antipodal decomposition for the unpolarized light is arbitrary [i.e., the vector  $\mathbf{v}(\chi', \psi')$  is not unique].

Apart from the nonunique decomposability of mixtures, it is important to appreciate that this geometric approach is appropriate to refine the following theorems (with reference to Fig. 3.2.1) discussed in Section 3.1:

1. Any quasimonochromatic light may be regarded as the sum of the completely unpolarized state and a completely polarized wave that are independent of each other.
2. The center of the Poincaré sphere has a unique status—it can be decomposed into infinitely many pairs of orthogonal states, for there are infinitely many diameters through it. This is reminiscent to the fact that the center, corresponding to the density matrix  $\mathbf{D}_{2u}$ , has a doubly degenerate eigenvalue. In other words, it has a spherical symmetry. To visualize such a decomposition, we notice that a wave of natural light of intensity  $\langle S_0 \rangle$  is equivalent to two independent circularly polarized waves, one left-handed ( $\circlearrowleft$ ), the other right-handed ( $\circlearrowright$ ), each of intensity  $\langle S_0 \rangle/2$ , representing the south and north poles of  $\Sigma_2^1$ .

It is also worth commenting that instead of choosing the normalized Stokes parameters  $\langle S_k \rangle / \langle S_0 \rangle = \langle \sigma_k \rangle$  as coordinate axes, one can choose the instantaneous Stokes parameters  $S_k / S_0 \equiv s_k$ . In that representation, the trajectory of the endpoint of

$$\mathbf{u} = \begin{bmatrix} \cos(2\chi) & \cos(2\psi) \\ \cos(2\chi) & \sin(2\psi) \\ \sin(2\chi) \end{bmatrix}$$

is an interesting graphical representation of the evolution of the state of polarization as light fluctuates. For instance, unpolarized light can be viewed as an erratic random walk on  $\Sigma_1^2$ , described by a uniform PDF over  $\Sigma_1^2$ . At the

other extreme, the PDF for a completely polarized state takes the form of a Dirac delta function, represented by a point on  $\Sigma_1^2$ . Thus an advantage offered by the description of partially polarized light in terms of  $s_k$  parameters is that it appears as a two-dimensional problem unlike the three-dimensional case required by the use of  $\langle \sigma_k \rangle$  parameters. A discussion of this topic is postponed until Section 3.3.3.3.

A nice feature of the Poincaré sphere representation is that it provides us with a natural coordinate system in which to analyze the effect of polarization transformation of an incoming wavefield through a linear optical instrument. We shall return to this problem in Parts 4 and 5, when we are better equipped to discuss it, but for the present we may remark that this claim is based on the homomorphism of the group of rotations in three dimensions  $SO(3)$  and the unimodular unitary group  $SU(2)$ .

### 3.2.2. CONVEXITY STRUCTURE OF SETS OF POLARIZATION STATES

This section contains an algebraic specification of the convex set of polarization states. We begin by collecting a few standard results. As we said above, the polarization states are in bijective correspondence with the density operators  $\mathbf{D}$  in the associated complex  $H$  Hilbert space of finite dimension. The set of all density matrices forms a convex set. If  $\mathbf{D}_j, j \in \langle 1, n \rangle$  are density operators in  $H$  and  $p_j$  are real positive numbers, such that

$$\sum_{j=1}^n p_j = 1$$

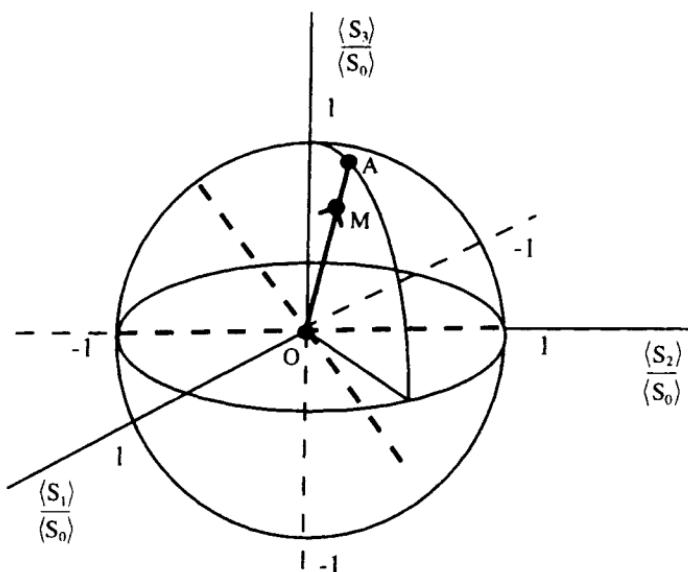
the convex combination

$$\sum_{j=1}^n p_j \mathbf{D}_j$$

is still a density operator. The expression is rarely unique. Such convex sets are stratified by rank. We define the rank of  $\mathbf{D}$  as the smallest number  $n$  of the  $\mathbf{D}_j$  needed, that is, the minimal number of extremal elements of which it is a convex combination. An element of a convex set is extremal if it is not a convex combination of two other elements. Thus the extremal states have rank one. The rank allows us to divide the set of density operators into strata that are subsets of equal rank. In a convex set, the set of elements of maximal rank has maximal dimension and forms the interior of the set. The interior is surrounded by a hierarchy of hypersurfaces of all lower ranks down to one. In the following the convexity structure of sets of polarization states is used to position pure states and mixed states relative to each other.

In the following the ideas mentioned above are developed. For plane waves described by the density  $\mathbf{D}_2$  matrix, we have  $n = 2$ . Thus, the convex set of

states has two strata of ranks 1 and 2 and respectively, dimensions 2 and 3. The pure states (of rank 1) lie on the surface of the Poincaré sphere, while mixed states (of rank 2) are the points inside the ball. When a polarization state cannot be written as a convex combination of other states, it represents a pure state (totally polarized light). The hallmark of pure states is that they are the extremal states of the convex set  $H$ . Within the geometric picture of the Poincaré sphere, they are localized at the surface of the unit ball  $\Sigma_1^3$ . The states that are convex combinations of others are called *mixtures* or *mixed states*; they describe partially polarized light and are localized inside  $\Sigma_1^3$ . As before, this parametrization also affords an easy representation in Stokes space, having a clear physical meaning. We have seen previously that  $D_2$  admits infinitely many decompositions into pure states. This is a consequence of the Krein–Millman theorem, which states that a compact convex set is completely determined by its extreme points. According to the Krein–Millman theorem, a compact convex set is completely determined by its extreme points; then any member of the set may be expressed as a linear combination of pure states. Moreover, the center of the unit ball  $\Sigma_1^3$  can be decomposed into an infinite number of pairs of orthogonal states corresponding to the many diameters through it (represented by dashed lines in Fig. 3.2.3). This unique state describes totally unpolarized light. Let us add a remark. In this representation the degree of polarization of a mixed state is given by the distance of point  $M$



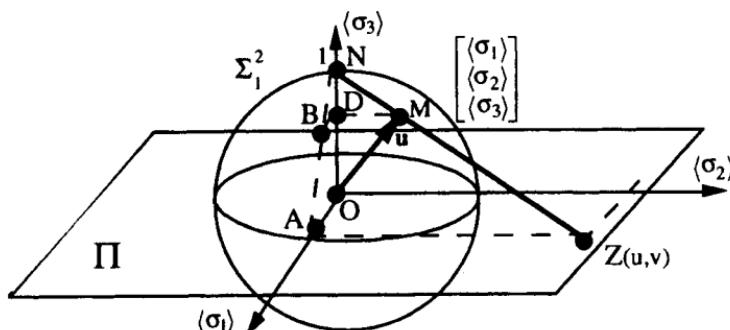
**Figure 3.2.3.** Schematic illustration of the convexity property of the set of polarization states on the unit ball  $\Sigma_1^3$ . Pure states correspond to surface points (e.g.,  $A$ ) and mixed states to interior points (e.g.,  $M$ ). The partially polarized state, represented by point  $M$ , is represented by the vector  $OM$  whose length is the degree of polarization  $P$ .

from the center of the sphere,  $|\mathbf{OM}| = P$ , where  $|\mathbf{OM}|$  indicates the vector norm of  $\mathbf{OM}$ . This expression is important not only because it defines geometrically the degree of polarization but also because it laid a conceptual foundation for analyzing the change in the degree of polarization of a wave when it interacts with matter. In Section 4.1 we shall have occasion to discuss this point in detail.

### 3.2.3. STEREOGRAPHIC PROJECTION PLANE OF POINCARÉ SPHERE

A second geometric construction, due to Poincaré, provides a means of visualizing the state of polarization of pure states of polarization. We have seen in Section 3.1.4 that, instead of using the variables  $\psi$  and  $\chi$  (or  $\alpha$  and  $\theta$ ), we can parametrize a pure state by a complex polarization ratio. This transformation constitutes the stereographic representation of the Poincaré sphere. This may be explained as a direct consequence of the algebraic structure of the underlying  $SU(2)$  group. This section is intended to show how elements of the unimodular unitary group  $SU(2)$  can be interpreted in terms of complex coordinates of points in a stereographic projection plane and by so doing establishes a one-to-one correspondence between the points of the plane  $\Pi$  and of all the points of the Poincaré sphere  $\Sigma_2^1$  except one, namely, the pole of projection. The use of complex coordinates can occasionally be advantageous, especially when studying the evolution of pure states propagating in non-depolarizing anisotropic media [9].

The principle of stereographic projection is illustrated in Fig. 3.2.4, where  $\Sigma_2^1$  has its center at the origin of the Cartesian coordinate system  $(\langle \sigma_1 \rangle, \langle \sigma_2 \rangle, \langle \sigma_3 \rangle)$ .



**Figure 3.2.4.** Stereographic projection through the north pole  $N$ , of the point  $M$  lying on the sphere  $\Sigma_2^1$ , onto the equatorial plane  $\Pi$ . This projection illustrates the relationship between the  $SU(2)$  Poincaré vector  $\mathbf{u}$  and the complex coordinates of point  $Z$ . From the similar triangles  $NOA$  and  $NDB$ , one obtains the coordinate  $u$  of  $Z$ . The value of  $v$  follows from an analogous construction. Thus we obtain the point  $Z$  with coordinates  $(u, v)$ .

$\langle \sigma_3 \rangle$ ). Stereographic coordinates are defined by the projection from the north pole  $N = [0, 0, 1]^T$  onto the equatorial plane.

Mathematically we construct the plane  $\Pi$  such that it coincides with the equator of  $\Sigma_2^1$ , and we suppose that each point of  $\Pi$  is joined to the north pole  $N$  by a straight line. For each point  $Z$  of  $\Pi$ , we can assign two coordinates  $(u, v)$  to every point  $M$  on  $\Sigma_2^1$ , with exception of the point  $N$ . The pole  $N$  of projection does not correspond to any of the complex numbers thus defined. This singularity is called the *point at infinity*.

The coordinates assigned in this way to a point

$$M(\psi, \chi) = \begin{bmatrix} \langle \sigma_1 \rangle \\ \langle \sigma_2 \rangle \\ \langle \sigma_3 \rangle \end{bmatrix}$$

of the sphere  $\Sigma_2^1$  are as follows. The line  $NM$  and the plane  $\Pi$  intersect at

$$Z = \begin{bmatrix} u \\ v \\ 0 \end{bmatrix}$$

Using the similar triangles  $NDB$  and  $NOA$  to scale the projection (see Fig. 3.2.4), we find that the stereographic coordinates of  $M$  are  $u = \langle \sigma_1 \rangle / (1 - \langle \sigma_3 \rangle)$  and  $v = \langle \sigma_2 \rangle / (1 - \langle \sigma_3 \rangle)$ . We may also write the stereographic coordinates  $(u, v)$  in terms of the polarization angles  $\chi$  and  $\psi$  [9]. We have

$$u = \cotan\left(\chi - \frac{\pi}{4}\right) \cos(2\psi) \quad (3.2.2a)$$

$$v = \cotan\left(\chi - \frac{\pi}{4}\right) \sin(2\psi) \quad (3.2.2b)$$

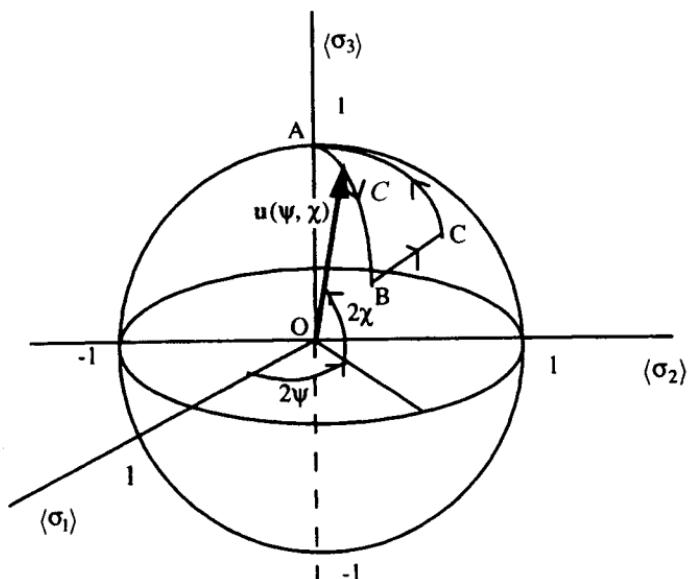
The stereographic projection of the Poincaré sphere onto plane  $\Pi$  has a clear geometric interpretation in that it allows the mapping of points on  $\Sigma_2^1$  to complex polarization ratios  $Z_{lr} = u + iv = (\langle \sigma_1 \rangle + i\langle \sigma_2 \rangle) / (1 - \langle \sigma_3 \rangle)$ . This property can be proved by substituting Eqs. (3.2.2a, b) in  $Z_{lr} = u + iv$ . We find  $Z_{lr} = \cotan[\chi - (\pi/4)] \exp(2i\psi)$ , which is recognized as the complex polarization ratio  $Z_{lr}$  given explicitly by placing Eq. (3.1.57) in Eq. (3.1.52). Thus, each point on the surface of  $\Sigma_2^1$ , except  $N$ , has assigned to it the point  $Z$  in the complex plane.

### 3.2.4. GEOMETRIC PHASE AND POLARIZATION CYCLES

In this section we have a twofold purpose. The primary purpose is to discuss the concept of geometric phase (also termed the *Berry–Pancharatnam* phase) for polarization circuits of light and to present a proof of an remarkable

theorem due to Pancharatnam concerning the phase change of a light beam whose polarization state is made to trace out a closed circuit on the Poincaré sphere: Pancharatnam's theorem guarantees that if an arbitrary pure state of polarization trace out a closed curve  $C$  on  $\Sigma_1^2$ , the polarization state acquires a geometric phase equal to half the solid angle subtended by  $C$  at the origin of  $\Sigma_1^2$ . The secondary purpose is to discuss this prediction vs experiment [30].

Before proceeding, we should inject an historical note. It has been about 40 years since the geometric phase concept was predicted and observed by Pancharatnam of Bangalore. Pancharatnam was investigating how two beams of light, in nonorthogonal states of polarization, could have the same phase. For that purpose, he considered the intensity of the wave obtained by linear superposition of the two beams. The two beams are defined as being in phase when the intensity is maximum. Now if we consider that the two beams can represent successive states of polarization as the light propagates, Pancharatnam's rule enables us to define how a beam can preserve its phase while its polarization state is altered. This rule is bijective but nontransitive. Pancharatnam then discovered the "half the solid angle" formula quoted above for the phase change of the polarization of a light beam tracing out a closed circuit on the Poincaré sphere, such as the path  $A \rightarrow B \rightarrow C \rightarrow A$  in Fig. 3.2.5. For simplicity, the closed circuit considered here is the spherical (Euler) triangle  $ABC$ . More recently, Berry considered a different situation by considering light with a fixed state of polarization and changing direction [31]. He pointed out



**Figure 3.2.5.** The phase change for a closed circuit  $C$  of polarization states is half the solid angle subtended by the circuit at the center of  $\Sigma_1^2$ . We consider the special situation of the circuit  $C$  made of the spherical triangle with one vertex at the north pole of  $\Sigma_1^2$ .

that the angle of rotation of the direction of polarized light after it has traveled along a fiber coiled into a helix is equal to the solid angle through which the fiber tangent has turned. The reader may wish to consult the pioneering papers of Pancharatnam and Berry on this subject. Manifestations of the geometric phase appeared in many fields of physics ranging from magnetic resonance [32] to neutron interferometry [33].

With these remarks, we now outline an elegant derivation of Pancharatnam's theorem due to Aravind [34]. An alternative method based on differential geometry is due to De Vito and Leviero [35]. First I need to explain the central idea. To get a start, we consider the situation described in the 1956 seminal paper of Pancharatnam [30]. If we consider an arbitrary pure state of polarization that traces out a geodesic triangle  $ABC$  on  $\Sigma_1^2$ , what is the phase difference between the initial and the final states? With reference to Fig. 3.2.5, consider the closed polarization circuit  $A \rightarrow B \rightarrow C \rightarrow A$  on  $\Sigma_1^2$ .

It is first instructive to derive an important result concerning the superposition of pure states. We define, in the usual way (see Section 3.1), the complex Jones vectors of two waves as

$$\mathbf{E}_a(\alpha_a, \theta_a) = \begin{bmatrix} \cos(\alpha_a) \\ \sin(\alpha_a) \exp(i\theta_a) \end{bmatrix}$$

and

$$\mathbf{E}_b(\alpha_b, \theta_b) = \exp(-i\delta) \begin{bmatrix} \cos(\alpha_b) \\ \sin(\alpha_b) \exp(i\theta_b) \end{bmatrix}$$

where  $\delta$  represents a relative phase between the two states. If we make the two waves interfere, we can write the intensity of the resulting wave is proportional to

$$|\mathbf{E}_a + \mathbf{E}_b|^2 = 2 + 2\text{Re}(\mathbf{E}_a^* \cdot \mathbf{E}_b) \quad (3.2.3)$$

It is a consequence of Eq. (3.2.5) that the maxima of intensity (according to the definition of Pancharatnam,  $\mathbf{E}_a$  and  $\mathbf{E}_b$  are in phase) must satisfy

$$\text{Re}(\mathbf{E}_a^* \cdot \mathbf{E}_b) > 0 \quad (3.2.4a)$$

$$\text{Im}(\mathbf{E}_a^* \cdot \mathbf{E}_b) = 0 \quad (3.2.4b)$$

Incidentally, note that Eqs. (3.2.4a, b) would not be valid if we were to consider the two Jones vectors  $\mathbf{E}_a$  and  $\mathbf{E}_b$  to form an orthonormal pair of vectors, as in Eq. (3.1.47). Since orthogonal states do not interfere, the phase difference between  $\mathbf{E}_a$  and  $\mathbf{E}_b$  remains undefined. The explicit (unique) solution for  $\delta$  is obtained by substituting the complex two-dimensional column vectors  $\mathbf{E}_a$  and  $\mathbf{E}_b$  into Eqs. (3.2.4a, b). We continue our discussion of the three sequential changes in polarization, from this point of view.

Next we consider three nonorthogonal pure states  $A$ ,  $B$  and  $C$ , which are such that  $A$  is in phase with  $B$  and  $B$  is in phase with  $C$ . Pancharatnam's contribution was to show that  $C$  is not in phase with  $A$  but with a state  $A'$  that differs in phase from  $A$  by an amount equal to half the area of the geodesic triangle  $ABC$ . Without loss of generality we consider the special situation of the circuit  $C$  made of the spherical triangle with one vertex at the north pole of  $\Sigma_1^2$  (Fig. 3.2.5). We may therefore write the three complex Jones vectors as

$$\mathbf{E}_A = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} \quad (3.2.5a)$$

$$\mathbf{E}_B = \exp(-i\delta_B) \begin{bmatrix} \cos(\alpha_B) \\ \sin(\alpha_B) \exp(i\theta_B) \end{bmatrix} \quad (3.2.5b)$$

$$\mathbf{E}_C = \exp(-i(\delta_C + \delta_B)) \begin{bmatrix} \cos(\alpha_C) \\ \sin(\alpha_C) \exp(i\theta_C) \end{bmatrix} \quad (3.2.5c)$$

where  $\delta_B$  (resp.  $\delta_C$ ) represents the relative phase between states  $B$  and  $A$  (resp.  $C$  and  $B$ ). We also assume that the state  $A'$  can be expressed in the form

$$\mathbf{E}_{A'} = \exp(-i\delta_{A'}) \mathbf{E}_A \quad (3.2.6)$$

where  $\delta_{A'}$  represents the relative phase between  $A'$  and  $A$ . We now invoke the condition Eqs. (3.2.4a, b) to calculate  $\delta_{A'}$ ,  $\delta_B$  and  $\delta_C$ . Equations (3.2.4a, b) imply that  $\delta_B = 0$  and  $\delta_{A'} = \delta_C$ . Next placing Eqs. (3.2.5b, c) in Eqs. (3.2.4a, b), we obtain

$$[\cos(\alpha_B) \cos(\alpha_C) + \cos(\theta_C) \sin(\alpha_B) \sin(\alpha_C)] \cos(\delta_C)$$

$$+ \sin(\theta_C) \sin(\alpha_B) \sin(\alpha_C) \cos(\delta_C) > 0 \quad (3.2.7a)$$

$$- \cos(\alpha_B) \cos(\alpha_C) \sin(\delta_C) + \sin(\alpha_B) \sin(\alpha_C) \sin(\theta_C - \delta_C) = 0 \quad (3.2.7b)$$

It follows directly from Eq. (3.2.7b) that we can write

$$\tan(\delta_{A'}) = \frac{\sin(\alpha_B) \sin(\alpha_C) \sin(\theta_C)}{[\cos(\alpha_B) \cos(\alpha_C) + \cos(\theta_C) \sin(\alpha_B) \sin(\alpha_C)]} \quad (3.2.8)$$

Following Aravind, we may use the formula for the area  $\Omega$  of the spherical triangle  $ABC$  on  $\Sigma_1^2$  (see details in Ref. 34)

$$\tan\left(\frac{\Omega}{2}\right) = \frac{|\mathbf{OA} \cdot \mathbf{OB} \times \mathbf{OC}|}{1 + \mathbf{OA} \cdot \mathbf{OB} + \mathbf{OB} \cdot \mathbf{OC} + \mathbf{OC} \cdot \mathbf{OA}} \quad (3.2.9)$$

In the situation envisaged (Fig. 3.2.5),

$$\mathbf{OA} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{OB} = \begin{bmatrix} \sin(2\alpha_B) \\ 0 \\ \cos(2\alpha_B) \end{bmatrix} \quad \mathbf{OC} = \begin{bmatrix} \sin(2\alpha_C) \cos(\theta_C) \\ \sin(2\alpha_C) \sin(\theta_C) \\ \cos(2\alpha_C) \end{bmatrix}$$

so that Eq. (3.2.10) reduces to

$$\tan\left(\frac{\Omega}{2}\right) = \frac{\sin(2\alpha_B) \sin(2\alpha_C) \sin(\theta_C)}{[(1 + \cos(2\alpha_B))(1 + \cos(2\alpha_C)) + \cos(\theta_C) \sin(2\alpha_B) \sin(2\alpha_C)]} \quad (3.2.10)$$

We leave to the reader the task to show that the right-hand sides of Eqs. (3.2.8) and (3.2.10) are equal. Consequently we arrive at the result

$$\delta_A = \frac{\Omega}{2} \quad (3.2.11)$$

Formula (3.2.11) is the primary result of this section and is precisely the result found by Pancharatnam. This asserts that if a pure state of polarization is made to trace out a closed circuit  $C$  on the Poincaré sphere, the phases of the final and initial states differ by half the solid angle subtended by  $C$  at the origin of  $\Sigma_1^2$ . Remarkably, Pancharatnam was also able to corroborate his theory experimentally [30].

A discussion of some experimental results is not without interest. An experimental detection of the geometric phase involves splitting a light beam into two parts, one of which has its state of polarization cycled through a closed circuit on  $\Sigma_1^2$  and the other of which has not undergone any change. Then interference of the beams reveals the geometric phase. People have exercised considerable ingenuity for measuring the geometric phase associated with a cycling of the wave polarization. The earliest confirmations of Pancharatnam's theorem were provided by the observations of Bhandari and Samuel [36] and Chyba et al. [37] using laser interferometry.

### 3.2.5. ANALOGY BETWEEN POLARIZATION AND TWO-LEVEL SYSTEMS: BLOCH EQUATION

Often in science, fruitful results come from combining two seemingly unrelated ideas into one. Here we shall indicate a useful geometric interpretation of polarization states in Poincaré space based on a formal analogy with two-level systems. In this section we consider the analogy between the equation of evolution of the Poincaré vector  $\mathbf{u}$  and the spin vector formalism of Bloch developed for magnetic resonance. Every physicist has tackled the quantum time evolution of two-level systems, such as effective spin- $\frac{1}{2}$  particle in a

magnetic field [21]. Several features of the formalism appropriate for two-level systems make it especially easy to investigate the issue of polarization evolution, such as light propagating through an optically anisotropic medium. This observation acquires more substance when one realizes that the two-level approach provides us with a geometric overview of the evolution of pure states.

The physical idea is the following. By way of background, let us recall some basic results of quantum mechanics concerning dynamical systems. The standard quantum-mechanical description is based on the complex Hilbert space. For a two-level system, this Hilbert space is complex two-dimensional and the group  $SU(2)$  plays a major role in the dynamics of these systems [38]. We start with the Bloch equation for the density matrix in the interaction representation. This has the form

$$i\hbar \dot{\mathbf{D}} = [\mathbf{H}, \mathbf{D}] \quad (3.2.12)$$

where  $\mathbf{H}$  is the Hamiltonian operator and the density operator  $\mathbf{D}$  specifies the time evolution of the field statistics. The overdot denotes a partial derivative with respect to time. For two-level systems, we may conveniently construct the  $2 \times 2$  Hamiltonian matrix by means of the generalized form

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} = \sum_{j=0}^3 H_j \boldsymbol{\sigma}_j \quad (3.2.13)$$

where  $H_{11}$  and  $H_{22}$  are real and  $H_{21} = H_{12}^*$ . Again restricting our consideration to a two-level system, it is useful to express the density matrix as [21]

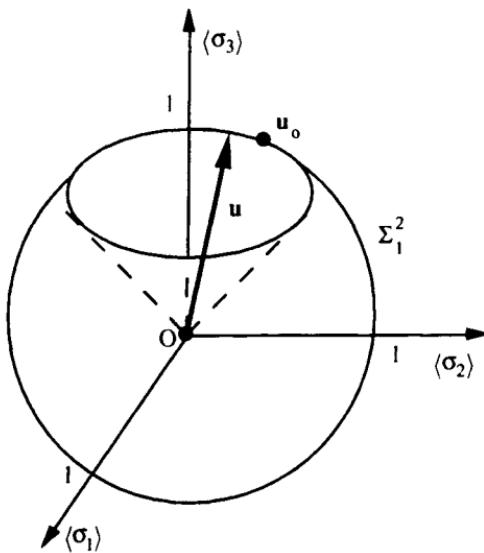
$$\mathbf{D} = \frac{1}{2} \left( \boldsymbol{\sigma}_0 + \sum_{j=3}^3 \langle \boldsymbol{\sigma}_k \rangle \boldsymbol{\sigma}_k \right) = \frac{1}{2} (\boldsymbol{\sigma}_0 + \mathbf{u} \cdot \boldsymbol{\sigma}) \quad (3.2.14)$$

which is the analog of Eq. (3.1.133). Next, substitution of Eqs. (3.2.13) and (3.2.14) into Eq. (3.2.12) yields the following equation for the three-dimensional vector

$$\mathbf{u} = \begin{bmatrix} \langle \boldsymbol{\sigma}_1 \rangle \\ \langle \boldsymbol{\sigma}_2 \rangle \\ \langle \boldsymbol{\sigma}_3 \rangle \end{bmatrix}; \dot{\mathbf{u}} = \boldsymbol{\Omega} \times \mathbf{u} \quad (3.2.15)$$

where the torque vector may be expressed as  $\boldsymbol{\Omega} \equiv 2\mathbf{H}/\hbar$ . Equation (3.2.15) is known as the *equation of motion for the two-level system*.

Thus we have established a formal link between spin dynamics and the evolution of the Poincaré vector. The analog of time in Eq. (3.2.15) is the spatial coordinate of the Poincaré vector. The geometric interpretation of Eq. (3.2.15) is that at each instant of time,  $\mathbf{u}(t)$  undergoes an infinitesimal rotation around the axis given by  $\mathbf{H}(t)$ , with angular velocity  $2|\mathbf{H}|/\hbar$ . Consequently, if  $\mathbf{H}$  is time-independent, Eq. (3.2.17) is the equation of motion of  $\mathbf{u}$  around  $\mathbf{H}$ ,



**Figure 3.2.6.** Schematic illustration of the precession of the Poincaré vector  $\mathbf{u}$  on the unit sphere. As time evolves, the tip of  $\mathbf{u}$  traces out a curve on  $\Sigma_1^2$ . The orbit of motion is closed and periodic starting from an initial state  $\mathbf{u}_0$ .

which is familiar from nuclear magnetic resonance, that is, precession of a spin- $\frac{1}{2}$  around the direction of a constant magnetic field. Physically this procedure permits visualization of the dynamics of the system in geometric terms. It is straightforward to show, on taking the dot product of Eq. (3.2.15) with  $\mathbf{u}$  and  $\boldsymbol{\Omega}$ , that  $\mathbf{u} \cdot \mathbf{u}$  and  $\boldsymbol{\Omega} \cdot \mathbf{u}$  are both invariants of the motion. The former implies that the motion takes place on  $\Sigma_1^2$ ; a pure state will remain pure. The latter implies that the motion is confined to a plane. Consequently, the orbit of motion is given by the intersection of the two preceding surfaces and is a circle lying on the Poincaré sphere  $\Sigma_1^2$ . Figure 3.2.6 displays a typical orbit of motion.

It is a simple exercise to prove that, under a stereographic projection, circles on  $\Sigma_1^2$  map onto circles in the complex plane. Therefore, the  $SU(2)$  orbit in the complex plane is also a circle.

Up to this point all we have shown is that the use of the Bloch equation is useful in translating polarization results into a somewhat equivalent mathematical form in terms of an effective spin- $\frac{1}{2}$  system. This may be cosmetically pleasing, but the analogy has much more powerful properties. A discussion of polarization beats and polarization echo, an analog of spin echo, using this analysis has been given by Zapasskii and Kozlov allowing possible applications in optical devices [39].

This concludes our general geometric discussion of partial polarization. The utility and the power of the Poincaré sphere representation will be illustrated in Part 4.

### SECTION 3.3

## Statistics of the Radiation Field

In many problems, we are not able of indicating the explicit form of the multivariate joint PDF  $p_N$  of Eq. (3.1.71) describing the statistical behavior of the electric field. Everything knowable about the electric field can, in principle, be calculated from this infinite set of density functions. Fortunately, Gaussian light, which interests us in many applications, is an exception, and we devote this chapter to this aspect of the problem. One of the most important properties of a complex Gaussian random process is that all the moments may be expressed in terms of the second-order ones. If the stochastic process is non-Gaussian, then the Gaussian approximation may still be a good one provided corrections due to higher-order correlations are small. Thus it is sufficient to limit the study of the first two moments of the stochastic field. Moreover, there are two other fundamental reasons why the discussion of Gaussian radiation fields has extensive practical application in polarization optics: (1) any linear combination of linearly transformed Gaussian processes is also a Gaussian process and (2) the central-limit theorem states that, under rather general conditions, the PDF of the sum of  $N$  independent random variables tends to a Gaussian PDF as  $N$  approaches infinity [4–6]. Light from a stationary thermal source, such as an incandescent bulb, originates almost entirely in spontaneous emission from a collection of excited atoms or molecules. Such radiation is a complicated system with a very large number of independent contributions and can be properly modeled, according the central-limit theorem, as a Gaussian random process.

The subject that we shall address in this chapter is the formulation of a rigorous and general treatment of the second-order statistical description of a stationary plane optical field. This chapter is laid out as follows. We start with the statistical analysis of amplitude and phases and statistics of the polarization ratio. Since these parameters are not observable at optical frequencies, this approach might be regarded as an instructive, but academic, exercise. Indeed, the reader with little taste for the subject can skip these sections entirely with little loss in understanding what follows, referring back to it on occasion for the elucidation of arcane terms. Attention is then turned to the statistics of the Stokes parameters. These quantities are of special importance because they are the observable quantities of the radiation field at optical frequencies. In Section 3.1 we described an unpolarized plane-wave radiation as an optical field that

could be represented by the diagonal density matrix  $\mathbf{D}_u = \frac{1}{2}\mathbf{\sigma}_0$ . In Section 3.3.4 we shall present a more accurate description of unpolarized light by appealing to certain symmetries that are well described by the Verdet–Stokes–Barakat conditions. Another point to emphasize at the start is that this section is more technical than the previous ones and some familiarity with the theory of random processes is required.

### 3.3.1. STATISTICS OF AMPLITUDES AND PHASES FOR A GAUSSIAN RANDOM WAVEFIELD

So far no assumption has been made about the probability density functions that govern the fluctuations of the optical field. It will now be assumed that the electromagnetic field components to be statistically homogeneous and zero-mean temporally stationary, at least in the wide sense, Gaussian random processes. We shall also assume that the light is adequately modeled as an ergodic process and that the waves form a narrow band about a large central frequency. Both components of the electric field obey the same statistics; their Hilbert transforms also obey the same statistics. Furthermore, the corresponding complex-valued analytic signals  $E_j(\mathbf{R},t)$  ( $j = 1, 2$ ) at a given position vector of a typical point in space and time  $t$  are also zero-mean temporally stationary at least in the wide sense. If  $E_j^{(r)}(\mathbf{R},t;T)$  is a Gaussian random process and  $E_j^{(i)}(\mathbf{R},t;T)$  is obtained from it by a linear transform, it follows that  $E_j^{(i)}(\mathbf{R},t;T)$  is also a Gaussian random process. Moreover by letting  $T \rightarrow \infty$  and taking averages over the ensemble of  $E_j^{(r)}(\mathbf{R},t;T)$ , it follows at once that the mean of  $E_j^{(i)}(\mathbf{R},t;T)$  is zero also.

We start from the expression of the joint bivariate PDF for the two fluctuating orthogonal components  $E_j(t)$  ( $j = 1, 2$ ) of the transverse electric field  $\mathbf{E}$  at a particular point in space with respect to a prescribed coordinate system

$$p(E_1, E_2) = \frac{\det(\mathbf{R})}{(\pi)^2} \exp\left(-\sum_{i,j=1}^2 E_i^* R_{ij} E_j\right) \quad (3.3.1)$$

where the Hermitian and positive definite  $2 \times 2$   $\mathbf{R}$  matrix has dimension of the inverse of an intensity [4–6]. In writing Eq. (3.3.1) we have taken the matrix  $\mathbf{R}$  such that  $\mathbf{R}^{-1} \equiv \Phi_2 = \langle S_0 \rangle \mathbf{D}_2$ . By definition,  $p(E_1, E_2)$  is a real-valued function of the complex Jones vector  $\mathbf{E}$ , which is nonnegative and normalized to unity.

We proceed as follows. To obtain information about the statistical properties of partially polarized light, we are interested in evaluating the moments of the electric field. For this purpose we write the components of the Jones vector  $E_j$  in terms of its amplitude  $a_j$  and phase  $\theta_j$ :  $E_j = a_j \exp(i\theta_j)$ . We will investigate the joint distribution of the  $a_j$  and  $\theta_j$  values as well as their marginal probability density functions [40]. The Jacobian of this transformation is  $a_1 a_2$ .

All these quantities can be evaluated in terms from the joint probability density function ( $a_1, a_2, \theta_1, \theta_2$  are taken at the same time)

$$p(a_1, a_2, \theta_1, \theta_2) = \frac{a_1 a_2}{(\pi)^2} \det(\mathbf{R}) \exp\left(-\sum_{i,j=1}^2 E_i^* R_{ij} E_j\right) \quad (3.3.2)$$

The joint distribution of amplitudes is given by phase integration of Eq. (3.3.2). It is convenient to define an additional phase by  $R_{12} = |R_{12}| \exp(i\phi_{12})$ . We may therefore rearrange Eq. (3.3.2) and write

$$\begin{aligned} p(a_1, a_2) &= \frac{a_1 a_2}{\pi^2} \det(\mathbf{R}) \exp(-(R_{11} a_1^2 + R_{22} a_2^2)) \\ &\times \int_0^{2\pi} \int_0^{2\pi} \exp(-2|R_{12}| a_1 a_2 \cos(\theta_1 - \theta_2 - \phi_{12})) d\theta_1 d\theta_2 \end{aligned} \quad (3.3.3)$$

Using the integral identity, we obtain

$$\frac{1}{2\pi} \int_0^{2\pi} \exp(a \cos \theta) d\theta = I_0(a) \quad (3.3.4)$$

where  $I_0$  is the modified Bessel function of zero order and first kind; we find that

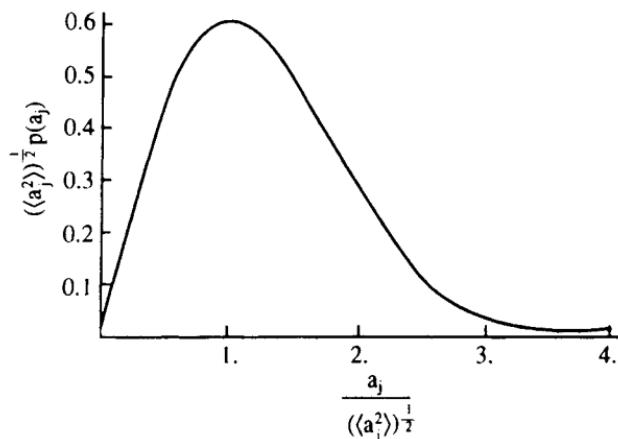
$$\begin{aligned} p(a_1, a_2) &= \begin{cases} \frac{16 a_1 a_2}{\langle S_0 \rangle^2 (1 - P^2)} \exp\left(\frac{-2(a_1^2 + a_2^2)}{\langle S_0 \rangle (1 - P^2)}\right) I_0\left(\frac{4P a_1 a_2}{\langle S_0 \rangle (1 - P^2)}\right) & \text{when } 0 < a_1, a_2 < \infty \\ 0 & \text{elsewhere} \end{cases} \\ & \quad (3.3.5) \end{aligned}$$

It is seen that, in general,  $p(a_1, a_2)$  cannot be written as a product of a function depending of  $a_1$  and a function depending on  $a_2$ . From Eq. (3.3.5) it follows that the marginal PDF of  $a_1$  and  $a_2$  are characteristic of a Rayleigh distribution

$$p(a_j) = \begin{cases} \frac{a_j}{\langle a_j^2 \rangle} \exp\left(-\frac{a_j^2}{2\langle a_j^2 \rangle}\right) & \text{when } 0 \leq a_j < \infty \\ 0 & \text{elsewhere} \end{cases} \quad (3.3.6)$$

The PDF is graphed in Fig. 3.3.1. The moments of  $a_j$  are given by

$$\langle a_j^{2l} \rangle = \Gamma(l + 1) \langle a_j^2 \rangle^l \quad (3.3.7)$$



**Figure 3.3.1.** The Rayleigh probability density function.

Now from Eq. (3.3.5) and provided that  $\text{Re}(\alpha) > -2$  and  $\text{Re}(\beta) > -2$ , we may derive the moments of order  $\alpha + \beta$ :

$$\langle a_1^\alpha a_2^\beta \rangle = \int_0^\infty \int_0^\infty a_1^\alpha a_2^\beta p(a_1, a_2) da_1 da_2 \quad (3.3.8)$$

Expanding the Bessel function  $I_0$  continued in  $p(a_1, a_2)$ , the result of the integration yields

$$\langle a_1^\alpha a_2^\beta \rangle = \frac{\Gamma(1 + \alpha/2)\Gamma(1 + \beta/2) \det(\mathbf{R})}{(R_{11})^{1+\alpha/2}(R_{22})^{1+\beta/2}} {}_2F_1\left(1 + \frac{\alpha}{2}, 1 + \frac{\beta}{2}; 1; \lambda^2\right) \quad (3.3.9)$$

where  ${}_2F_1$  is the usual hypergeometric function and  $\lambda \equiv |R_{12}|/(R_{11}R_{22})^{1/2} = |\gamma_{12}|$  is the modulus of the temporal degree of coherence [40]. As an illustration of Eq. (3.3.9), we consider integer values of  $\alpha$  and  $\beta$  (let  $\alpha = 2p$ ,  $\beta = 2q$ , for instance). Using Eq. (3.3.9), we have

$$\langle |E_1|^{2p} |E_2|^{2q} \rangle = \frac{\Gamma(1 + p)\Gamma(1 + q) \det(\mathbf{R})}{(R_{11})^{1+p}(R_{22})^{1+q}} {}_2F_1(1 + p, 1 + q, 1; \lambda^2) \quad (3.3.10)$$

which reduces to

$$\langle |E_1|^{2p} |E_2|^{2q} \rangle = \Gamma(p + 1)\Gamma(q + 1)(\langle |E_1|^2 \rangle)^p(\langle |E_2|^2 \rangle)^q \sum_{r=0}^q \binom{p}{r} \binom{q}{r} \lambda^{2r} \quad (3.3.11)$$

if we make use of the identity [41]

$${}_2F_1(1+p, 1+q, 1; x) = (1-x)^{-(p+q+1)} \sum_{r=0}^q \binom{p}{r} \binom{q}{r} x^r \quad (3.3.12)$$

From Eq. (3.3.11) and the usual properties of the complete elliptic integrals of the first and second kind (resp.  $K$  and  $E$ ), we may derive the covariance of  $a_1$  and  $a_2$ :

$$\begin{aligned} \text{cov}(a_1, a_2) &= \langle a_1 a_2 \rangle - \langle a_1 \rangle \langle a_2 \rangle \\ &= \frac{1}{2} (\langle |E_1|^2 \rangle \langle |E_2|^2 \rangle) \left[ 2E(\lambda) - (1-\lambda^2)K(\lambda) - \frac{\pi}{2} \right] \end{aligned} \quad (3.3.13)$$

where the following identities have been used [41]

$$(1-\lambda^2) {}_2F_1\left(\frac{3}{2}, \frac{3}{2}, 1; \lambda^2\right) = {}_2F_1\left(-\frac{1}{2}, -\frac{1}{2}, 1; \lambda^2\right) = \frac{2}{\pi} [2E(\lambda) - (1-\lambda^2)K(\lambda)] \quad (3.3.14)$$

$${}_2F_1\left(\frac{1}{2}, \frac{1}{2}, 1; \lambda\right) = \frac{2}{\pi} K(\lambda) \equiv \frac{2}{\pi} \int_0^{\pi/2} \Delta^{-1}(\alpha) d\alpha \quad (3.3.15a)$$

and

$${}_2F_1\left(-\frac{1}{2}, \frac{1}{2}, 1; \lambda\right) = \frac{2}{\pi} E(\lambda) \equiv \frac{2}{\pi} \int_0^{\pi/2} \Delta(\alpha) d\alpha \quad (3.3.15b)$$

with  $\Delta(\alpha) \equiv (1 - r \sin^2 \alpha)^{1/2}$ . Both  $K$  and  $E$  are extensively tabulated; see, for instance, Abramovitz and Stegun [41]. It is clear from Eq. (3.3.15b), with  $\alpha = 1$ ,  $\beta = -1$  and if use is also made of

$${}_2F_1(a, b, c; x) = (1-x)^{c-a-b} {}_2F_1(c-a, c-b, c; x) \quad (3.3.16)$$

that

$$\left\langle \frac{a_1}{a_2} \right\rangle = \frac{\pi \det(\mathbf{R})}{2(R_{11})^{3/2}(R_{22})^{1/2}} {}_2F_1\left(\frac{3}{2}, \frac{1}{2}, 1; \lambda\right) = \left( \frac{\langle |E_1|^2 \rangle}{\langle |E_2|^2 \rangle} \right) E(\lambda) \quad (3.3.17)$$

As a further result, if  $\alpha = \beta = -1$ , we find

$$\begin{aligned} \left\langle \frac{1}{a_1 a_2} \right\rangle &= \frac{\pi \det(\mathbf{R})}{(R_{11} R_{22})^{1/2}} {}_2F_1\left(\frac{1}{2}, \frac{1}{2}, 1; \lambda\right) \\ &= \frac{2}{(\langle |E_1|^2 \rangle \langle |E_2|^2 \rangle)^{1/2}} K(\lambda) \end{aligned} \quad (3.3.18)$$

by making use of the relation (3.3.15a).

A similar analysis holds for the statistical characterization of the phases. The joint probability density function of phases is given by integration over the complete amplitude range  $0 \leq a_j < \infty$  ( $j = 1, 2$ ):

$$p(\theta_1, \theta_2) = \frac{\det(\mathbf{R})}{\pi^2} \int_0^\infty \int_0^\infty a_1 a_2 \exp(-(R_{11}a_1^2 + R_{22}a_2^2)) \times \exp(-2|R_{12}|a_1 a_2 \cos(\theta_1 - \theta_2 - \phi_{12})) da_1 da_2 \quad (3.3.19)$$

The integral in Eq. (3.3.19) can be evaluated with the help of the transformations

$$\begin{aligned} a_1 &= \langle S_0 \rangle (1 - P^2)^{1/2} z^{1/2} \exp(z) \\ a_2 &= \langle S_0 \rangle (1 - P^2)^{1/2} z^{1/2} \exp(-z) \end{aligned} \quad (3.3.20)$$

We obtain

$$p(\theta_1, \theta_2) = \frac{1 - P^2}{\pi^2} \int_0^\infty z \exp(\mu z) \int_{-\infty}^\infty \exp(-z \cosh(\phi)) d\phi dz \quad (3.3.21)$$

where the new variables  $(z, \phi)$  are such  $0 \leq z < \infty$  and  $-\infty < \phi < \infty$ . Integration yields

$$p(\theta_1, \theta_2) = \frac{1 - P^2}{\pi^2} \int_0^\infty z \exp(\mu z) K_0(z) dz \quad |\mu| \leq 1 \quad (3.3.22)$$

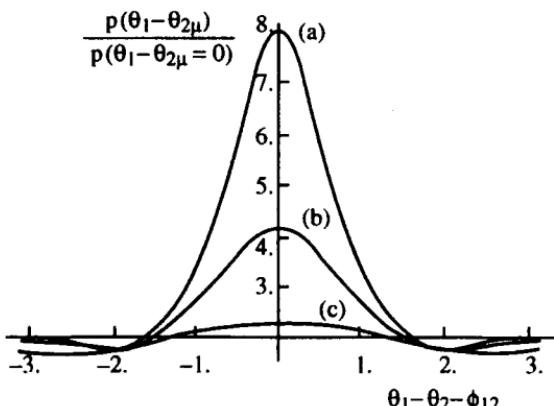
in which  $K_0$  is a modified Bessel function of the second kind

$$K_0(z) = \frac{1}{2} \int_{-\infty}^\infty \exp(-z \cosh(\phi)) d\phi \quad (z \geq 0) \quad (3.3.23)$$

and where  $\mu$  is defined as  $\mu = \lambda \cos(\theta_1 - \theta_2 - \phi_{12})$ . Evaluating Eq. (3.3.22) in closed form leads to the MacDonald-Bunimovitch relation [40, 42]

$$p(\theta_1, \theta_2) = \begin{cases} \frac{\det(\mathbf{R})}{8\pi^2 R_{11} R_{22}} (1 - \mu^2)^{-3/2} \left[ \mu \sin^{-1}(\mu) + \frac{\pi}{2} \mu + (1 - \mu^2)^{1/2} \right] & \text{if } -\pi \leq \theta_1, \theta_2 \leq \pi \\ 0 & \text{otherwise} \end{cases} \quad (3.3.24)$$

Figure 3.3.2 shows plots  $p(\theta_1, \theta_2, \mu)/p(\theta_1, \theta_2, \mu = 0)$  against  $\theta_1 - \theta_2 - \phi_{12}$  for various values of the parameter  $\lambda$ . Here we only note the following property of the joint probability density function of phases. With increasing  $\lambda$ , the shape



**Figure 3.3.2.** Plots of the reduced PDF  $p(\theta_1, \theta_2, \mu)/p(\theta_1, \theta_2, \mu = 0)$  of the phases (a)  $\lambda = 0.7$ , (b)  $\lambda = 0.5$ , and (c)  $\lambda = 0.1$ .

of the becomes more narrow, converging toward a  $\delta$  function at  $\lambda = 1$ . When  $\lambda = 0$ , the PDF is uniform on the primary interval  $[-\pi, \pi]$ .

An alternative representation of the joint PDF of phases can be constructed in the form

$$p(\theta_1, \theta_2) = \begin{cases} \frac{\det(\mathbf{R})}{8\pi^2 R_{11} R_{22}} \frac{\partial^2}{\partial \mu^2} (\cos^{-1}(\mu))^2 & \text{if } -\pi \leq \theta_1, \theta_2 \leq \pi \\ 0 & \text{elsewhere} \end{cases} \quad (3.3.25)$$

Since  $\theta_1$  and  $\theta_2$  appear only as  $\theta_1 - \theta_2$ , we may write  $p(\theta_1, \theta_2) = p(\theta_1 - \theta_2)$  as a function of the phase difference only. It is also worth noting that  $\theta_1$  and  $\theta_2$  have a rectangular PDF on the primary interval  $[-\pi, \pi]$

$$p(\theta_j) = \begin{cases} \frac{1}{2\pi} & \text{if } -\pi < \theta_j \leq \pi \\ 0 & \text{otherwise} \end{cases} \quad (3.3.26)$$

as any other possible density function would induce nonstationarity. A small additional point may be made about the PDF of  $\cos(\theta_j)$  and  $\sin(\theta_j)$  ( $j = 1, 2$ ). We arrive at the form

$$p_g = \begin{cases} \frac{1}{\pi(1 - g^2)^{1/2}} & \text{when } |g|^2 \leq 1 \\ 0 & \text{elsewhere} \end{cases} \quad (3.3.27)$$

where  $g = \cos(\theta_j)$  or  $g = \sin(\theta_j)$ . All odd moments of  $p_g$  are zero, while the first two even moments are  $\langle g^2 \rangle = \frac{1}{2}$  and  $\langle g^4 \rangle = \frac{3}{8}$ . Turning to the evaluation of the covariance of the phases, we find after simple but tedious manipulations [40]

$$\begin{aligned} \text{cov}(\theta_1, \theta_2) &= \langle \theta_1 \theta_2 \rangle - \langle \theta_1 \rangle \langle \theta_2 \rangle = 2 \sum_{m=1}^{\infty} \frac{1}{m^2} \lambda^m \frac{\Gamma^2(1 + 1/2m)}{\Gamma(1 + m)} \cos(m\phi_{12}) \\ &\quad \times {}_2F_1\left(\frac{m}{2}, \frac{m}{2}, m + 1; \lambda^2\right) \end{aligned} \quad (3.3.28)$$

It is instructive to specialize these results to the important cases of unpolarized light, (i.e.,  $\mathbf{R}$  is multiple of the  $2 \times 2$  unit matrix  $\sigma_0$ ), and completely polarized light. Consider first the unpolarized situation. The joint PDF, Eq. (3.3.3), reduces to a simple product of the marginal PDF  $p(a_j)$  and  $p(\theta)$ :

$$p(a_1, a_2, \theta) = p(a_1)p(a_2)p(\theta) \quad (3.3.29)$$

and implies that the  $a_j$  terms and  $\theta$  are independent random processes. Intuitively, this is what we would expect. At the other extreme case of completely polarized light, the joint PDF also factors in according to Eq. (3.3.29), again implying the statistical independence of  $a_1$ ,  $a_2$ , and  $\theta$ . It is important to note that the above results are independent of the particular form of the spectral density  $g(v)$  of the radiation, provided it has an arbitrarily narrow spectral extent.

### 3.3.2. STATISTICS OF THE COMPLEX POLARIZATION RATIO FOR A GAUSSIAN RANDOM WAVEFIELD

Our discussion of the complex polarization ratio in Section 3.1 was restricted to completely polarized light. For a partially polarized wavefield, the complex number  $Z_{12} \equiv E_2/E_1$  is a random process. Our formalism retains the general assumptions described in section 3.3.1. Given these assumptions, we can compute the average value of  $Z_{12}$  for a Gaussian field. From Eq. (3.3.1) and making use of

$$\begin{aligned} \int_0^{2\pi} \int_0^{2\pi} \exp(-2a_1 a_2 |R_{12}|) \cos(\theta_1 - \theta_2 - \phi_{12}) \cos(\theta_2 - \theta_1) d\theta_1 d\theta_2 \\ = -4\pi^2 I_1(2a_1 a_2 |R_{12}|) \cos(\phi_{12}) \end{aligned} \quad (3.3.30)$$

where  $I_1$  is the modified Bessel function of the first kind and of order 1, we

obtain the following expression for the average value of the complex variable  $Z_{12}$ :

$$\langle Z_{12} \rangle = -4 \exp(i\phi_{12}) \det(\mathbf{R}) \int_0^\infty \int_0^\infty a_1^2 \\ \times \exp(-(a_1^2 R_{11} + a_2^2 R_{22})) I_1(2a_1 a_2 |R_{12}|) da_1 da_2 \quad (3.3.31)$$

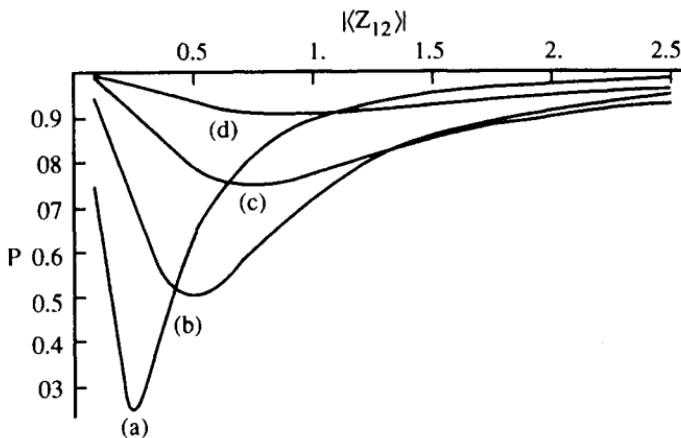
Integration yields

$$\begin{aligned} \langle Z_{12} \rangle &= \frac{-|R_{12}| \det(\mathbf{R})}{(R_{11})^2 R_{22}} \exp(i\phi_{12}) {}_2F_1(2, 1, 2; \lambda) \\ &= -\frac{R_{12}}{R_{11}} = \frac{\langle E_1 E_2^* \rangle}{\langle |E_1|^2 \rangle} = \frac{\langle S_2 \rangle - i\langle S_3 \rangle}{\langle S_0 \rangle + \langle S_1 \rangle} \end{aligned} \quad (3.3.32)$$

It can also be quickly established that

$$\frac{\text{Im}(\langle Z_{12} \rangle)}{\text{Re}(\langle Z_{12} \rangle)} = -\frac{\langle S_3 \rangle}{\langle S_2 \rangle} \quad (3.3.33)$$

For unpolarized light,  $\langle Z_{12} \rangle \equiv 0$ , whereas for completely polarized light,  $\langle Z_{12} \rangle = Z_{12}$ . We emphasize two important facts. First note that  $\langle Z_{12} \rangle$ , unlike  $P$  and  $\langle S_0 \rangle$ , is not invariant under the coordinate system. In other words, knowledge of  $\langle Z_{12} \rangle$  with respect to the 1–2 coordinate system is not sufficient



**Figure 3.3.3.** The degree of polarization as a function of the modulus of the average complex ratio for different values of the modulus of the degree of coherence  $|\gamma_{12}|$ : (a)  $|\gamma_{12}| = 0.90$ , (b)  $|\gamma_{12}| = 0.75$ , (c)  $|\gamma_{12}| = 0.50$ , (d)  $|\gamma_{12}| = 0.25$ .

to determine  $\langle Z_{uv} \rangle$  for any other  $u-v$  coordinate system [43]. Second,  $\langle Z_{12} \rangle$  does not uniquely determine the state of polarization, that is,  $\langle Z_{12} \rangle \equiv 0$ , for both unpolarized light and linearly polarized light ( $\langle S_2 \rangle = \langle S_3 \rangle = 0$ ). Of special interest is the expression of the degree of polarization in terms of the modulus of the degree of coherence and the modulus of  $\langle Z_{12} \rangle$ :

$$P = \left( 1 - 4(1 - |\gamma_{12}|^2) \left( \frac{|\gamma_{12}|}{|\langle Z_{12} \rangle|} + \frac{|\langle Z_{12} \rangle|}{|\gamma_{12}|} \right)^{-2} \right)^{1/2} \quad (3.3.34)$$

This equation implies that  $|\gamma_{12}| \leq P$ , as required. When  $|\gamma_{12}|$  is given,  $P$  reaches a minimum when  $|\langle Z_{12} \rangle| = |\gamma_{12}|$ ; this minimum is equal to  $|\gamma_{12}|$  itself; in this state  $P = |\langle Z_{12} \rangle| = |\gamma_{12}|$ , that is,  $\Phi_{11} = \Phi_{22}$ . Figure 3.3.3 is a plot of  $P$  against  $\langle Z_{12} \rangle$  for different values of  $|\gamma_{12}|$ .

### 3.3.3. STATISTICS OF STOKES PARAMETERS FOR A GAUSSIAN FIELD

Through the manipulations of the previous sections we have arrived at a detailed statistical characterization of the amplitudes and phases for a Gaussian random wavefield. Although the determination of this statistics is interesting, these parameters are not observable quantities at optical frequencies. Consequently it is necessary to develop and exploit the formalism appropriate for second-order statistics of the wavefield. The methods we are about to describe deal with the calculation of the probability density functions of the Stokes parameters and its variant, the normalized Stokes parameters and the time-integrated Stokes parameters. This section is devoted to the analysis of these functions and their polarization dependence.

#### 3.3.3.1. Probability Density Functions of Stokes Parameters

In developing the statistics of Stokes parameters, we assume in this section (as in Sections 3.3.1 and 3.3.2) that the quasimonochromatic wavefield is Gaussian distributed and stationary in time. Our goal is to calculate explicitly the PDF of the Stokes parameters  $S_0$  and  $S_1$ . We emphasize at the outset that the PDF of  $S_2$  will follow from that of  $S_1$  via a symmetry argument. It follows from Eq. (3.3.1) that the characteristic functions of  $S_0$  and  $S_1$  satisfy [44]

$$\begin{aligned} CS_0(u) &= \langle \exp(iuS_0) \rangle \\ &= \frac{\det(\mathbf{R})}{(2\pi)^2} \iint dE_1 dE_2 \exp\left(-\frac{1}{2} \sum_{i,j=1}^2 E_i^*(\mathbf{R}_{ij} - 2iu\delta_{ij})E_j\right) \end{aligned} \quad (3.3.35)$$

which can be also written as

$$\begin{aligned} CS_0(u) &= \det(\mathbf{R})(\det(R_{ij} - 2iu\delta_{ij}))^{-1} = \det(\mathbf{R}) \prod_{j=1}^2 (a_j - 2iu)^{-1} \\ &= \left(1 - \frac{2iu}{a_1}\right)^{-1} \left(1 - \frac{2iu}{a_2}\right)^{-1} \end{aligned} \quad (3.3.36)$$

Evaluation of the characteristic function of  $S_1$  follows the same technique as discussed above. The result is

$$CS_1(u) = \det(\mathbf{R}) (a_1 - 2iu)^{-1} (a_2 + 2iu)^{-1} = \left(1 - \frac{2iu}{a_1}\right)^{-1} \left(1 + \frac{2iu}{a_2}\right)^{-1} \quad (3.3.37)$$

where the  $a_j$  denote the eigenvalues of  $\mathbf{R}$  and  $\delta_{ij}$  is the Kronecker symbol. If there is any degeneracy (e.g.,  $P = 0$ ), the expressions have to be modified by taking the limiting value when  $a_1 \rightarrow a_2$ .

The PDF of  $S_j$  can be obtained by writing the inverse Fourier transform of the characteristic function  $CS_j$ . For  $S_0$ , we find

$$p(S_0) = \frac{\det(\mathbf{R})}{2\pi} \int \frac{\exp(-iuS_0)du}{(a_1 - 2iu)(a_2 - 2iu)} = -\frac{a_1 a_2}{8\pi} \int \frac{\exp(-iuS_0)du}{\left(u + i\frac{a_1}{2}\right)\left(u + i\frac{a_2}{2}\right)} \quad (3.3.38)$$

Since  $S_0 > 0$ , the integral converges and can be simply evaluated by going into the complex  $z$  plane. All the singularities of the integrand lie in the lower half-plane. Consequently, our contour of integration is the entire real line plus a semicircle in the lower half of the complex  $z$  plane. Application of Jordan's lemma shows that the semicircle does not contribute. Hence the integral over the real axis is equal to the sum of the residues. Since there are two simple poles at  $z = i(a_1/2), i(a_2/2)$ , we arrive at the relation

$$p(S_0) = \frac{a_1 a_2}{2(a_1 - a_2)} \left[ \exp\left(-\frac{a_1 S_0}{2}\right) - \exp\left(-\frac{a_2 S_0}{2}\right) \right] \quad (3.3.39)$$

with  $a_1 = u/[\langle S_0 \rangle(1 - P)]$  and  $a_2 = u/[\langle S_0 \rangle(1 + P)]$  ( $0 < a_2 < a_1$ ). Hence

$$p(S_0) = \frac{1}{P\langle S_0 \rangle} \left[ \exp\left(-\frac{2S_0}{\langle S_0 \rangle(1 + P)}\right) - \exp\left(-\frac{2S_0}{\langle S_0 \rangle(1 - P)}\right) \right] \quad \text{for } S_0 \geq 0 \quad (3.3.40)$$

and zero for  $S_0 < 0$ . Equation (3.3.40) contains several things worth mentioning. We can observe that  $p(S_0)$  depends only on the measurable parameters

$\langle S_0 \rangle$  and  $P$ . Now if  $P = 1$ , we find a negative exponential for  $p(S_0)$ :

$$p(S_0) = \frac{1}{\langle S_0 \rangle} \exp\left(-\frac{S_0}{\langle S_0 \rangle}\right) \quad \text{for } S_0 \geq 0 \quad (3.3.41)$$

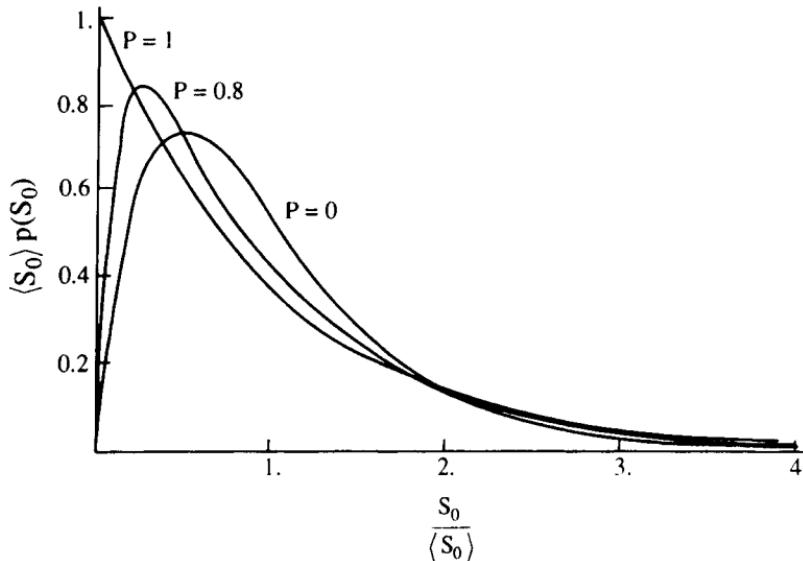
and zero for  $S_0 < 0$ . Numerical calculations of  $\langle S_0 \rangle p(S_0)$  versus  $S_0/\langle S_0 \rangle$  are shown in Fig. 3.3.4 for various values of the degree of polarization. As can be seen in this figure, all the curves start from the origin, except the one for  $P = 1$ . The corresponding result for unpolarized light is obtainable in analogous fashion:

$$p(S_0) = \frac{4S_0}{\langle S_0 \rangle^2} \exp\left(-\frac{2S_0}{\langle S_0 \rangle}\right) \quad \text{for } S_0 \geq 0 \quad (3.3.42)$$

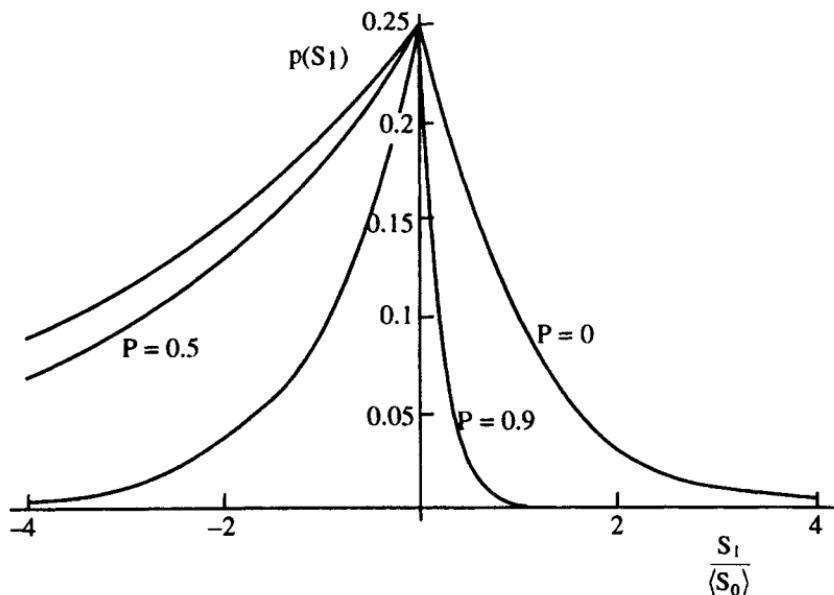
Equations (3.3.40)–(3.3.42) have the physical significance that unpolarized light has less probability of having small values of instantaneous intensity than does completely polarized light as illustrated in Fig. 3.3.4.

The PDF of  $S_1$  can be obtained in much the same manner. It yields

$$p(S_1) = \frac{\det(\mathbf{R})}{2\pi} \int \frac{\exp(-iuS_1)du}{(a_1 - 2iu)(a_2 + 2iu)} = \frac{a_1 a_2}{8\pi} \int \frac{\exp(-iuS_1)du}{\left(u + i\frac{a_1}{2}\right)\left(u - i\frac{a_2}{2}\right)} \quad (3.3.43)$$



**Figure 3.3.4.** The probability density function of the Stokes parameter  $S_0$  for three values of the degree of polarization.



**Figure 3.3.5.** The probability density function of the Stokes parameter  $S_1$  for two values of the degree of polarization.

The integration in Eq. (3.3.43) can be carried out in the same way as for Eq. (3.3.38). After an integration over a contour for which the integral over the boundaries is found to converge (note that we have to discriminate between  $S_1 > 0$  and  $S_1 < 0$ ), we arrive at the relation

$$p(S_1) = \frac{1}{4\langle S_0 \rangle} \exp\left(-\frac{|S_1|}{2\langle S_0 \rangle(1 - P \operatorname{sgn}(S_1))}\right) \quad (3.3.44)$$

where  $\operatorname{sgn}(u) \equiv u/|u|$ . Thus the PDF of  $S_1$  is asymmetric except for the special case of unpolarized light for which it takes a symmetric form. The result is shown plotted in Fig. 3.3.5. Note that the effect of increasing the degree of polarization is to force the peak of the PDF at  $S_1 = 0$  to increase. For completely polarized light,  $p(S_1)$  reduces to a Dirac delta function

$$\lim_{P \rightarrow 1} \langle S_1 \rangle p(S_1) = \delta(S_1) \quad (3.3.45)$$

In discussing Stokes parameters, we have already seen that  $S_2$  and  $S_3$  can be derived from  $S_1$  with a change of coordinates. These transformations are  $2 \times 2$  unitary matrices  $\mathbf{R}_S$  leaving the total intensity of the field  $\langle S_0 \rangle$  and the

degree of polarization  $P$  invariant:

$$S_1 \xrightarrow{\mathbf{R}_2} S_2 \quad \text{with} \quad \mathbf{R}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \quad (3.3.46a)$$

$$S_1 \xrightarrow{\mathbf{R}_3} S_3 \quad \text{with} \quad \mathbf{R}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \quad (3.3.46b)$$

The PDF of  $S_2$  and  $S_3$  can be immediately obtained applying this symmetry argument. While  $S_0$  obeys a PDF that is a difference of two negative exponentials, one finds in this way that the other three Stokes parameters are Laplacian variates [16, 44, 45]. To close this section, we add a note about the fact that few published experimental data on the subject have come to our attention except for the complementary problem of the spatial statistics of the Stokes parameters [46].

### 3.3.3.2. Cumulants of Stokes Parameters

So far we have considered only the PDF of the Stokes parameters. An alternative statistical characterization of these parameters involves the cumulants. The cumulants are defined as the coefficients of  $u^k/[\Gamma(k+1)]$  in a MacLaurin expansion of  $\ln(C(-iu))$ ,

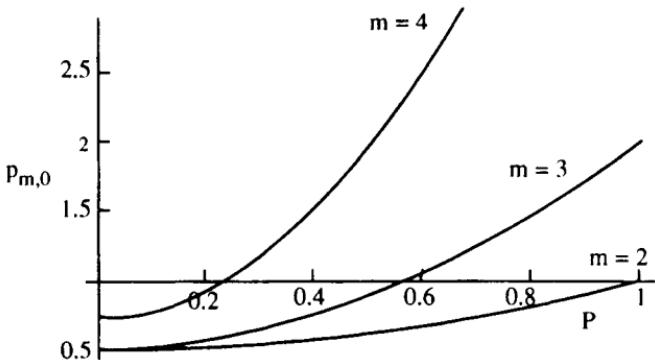
$$\ln\langle\exp(uX)\rangle = \sum_{k=1}^{\infty} \frac{u^k}{\Gamma(k+1)} \chi_k(X) \quad (3.3.47)$$

where  $\Gamma(k)$  is the Gamma function. For those familiar with the theory of random functions, we note that the first four cumulants have clear physical meanings. The first cumulant  $\chi_1(S_k) = \langle S_k \rangle$  is the mean of  $S_k$ . The second cumulant  $\chi_2(S_k) = \text{var}(S_k)$  defines the spreading of the probability density function of  $S_k$  around its mean. The third cumulant does not have a name but the dimensionless parameter  $\chi_3/(\chi_2)^{3/2}$  is often referred to as the *skewness factor* and serves as one indicator of the asymmetry of the  $\text{PDF}(S_k)$  around its mean. Similarly, the fourth cumulant does not have a name, but the flatness (also termed *kurtosis*) factor is defined as  $\chi_4/(\chi_2)^2$ . It is worth noting that these cumulants are also expressible in terms of moments.

From the very definition of cumulants, we find that [47]

$$\chi_m(S_k) = \Gamma(m) \text{ tr}((\mathbf{R}_k \Phi \mathbf{R}_k^{-1} \boldsymbol{\sigma}_1)^m) = \Gamma(m) \text{ tr}((\Phi \boldsymbol{\sigma}_k)^m) \quad (3.3.48)$$

if we make use of the relations  $\mathbf{R}_2^{-1} \boldsymbol{\sigma}_1 \mathbf{R}_2 = \boldsymbol{\sigma}_2$  and  $\mathbf{R}_3^{-1} \boldsymbol{\sigma}_1 \mathbf{R}_3 = \boldsymbol{\sigma}_3$ . The



**Figure 3.3.6.** Graph of the normalized cumulants  $p_{m,0}$  versus degree of polarization for  $m = 2, 3, 4$ .

normalized cumulants of order  $m$  of the Stokes parameters  $S_k$

$$p_{m,k} = \frac{\chi_m(S_k)}{(\chi_1(S_k))^m} \quad (3.3.49)$$

may be derived with the aid of Eq. (3.3.48). For this purpose, we need to compute the trace of the  $m$ th power of a  $2 \times 2$  matrix; this trace is simply the sum of the  $m$ th power of the eigenvalues  $\lambda_k^{(i)}$  ( $i = 1, 2$ ) of  $\Phi\sigma_k$ . Thus  $p_{m,k}$  has the form

$$p_{m,k} = \Gamma(m) \frac{\sum_{i=1}^2 (\lambda_k^{(i)})^m}{(\sum_{i=1}^2 \lambda_k^{(i)})^m} \quad (3.3.50)$$

where

$$\lambda_k^{(i)} = \lambda_k^{(\pm)} = \frac{\langle S_k \rangle}{2} \{1 \pm (1 + 4(1 - P^2)\langle \sigma_k \rangle^2)^{1/2}\} \quad \text{for } k \neq 0 \quad (3.3.51)$$

and

$$\lambda_0^{(i)} = \lambda_0^{(\pm)} = \frac{\langle S_0 \rangle}{2} \{1 \pm P\} \quad \text{for } k = 0 \quad (3.3.52)$$

Placing Eq. (3.3.52) into Eq. (3.3.50), we find that

$$p_{m,0} = \frac{\Gamma(m)}{2^m} [(1 + P)^m + (1 - P)^m] \quad (3.3.53)$$

In Fig. 3.3.6, we plot  $p_{m,0}$  versus  $P$  for three values of  $m$ .

Substitution of Eq. (3.3.51) into Eq. (3.3.50) gives

$$\begin{aligned} p_{m,k} = \frac{\Gamma(m)}{2^m} & [ \{1 + (1 + 4(1 - P^2)\langle\sigma_k\rangle^2)^{1/2}\}^m \\ & + \{1 - (1 + 4(1 - P^2)\langle\sigma_k\rangle^2)^{1/2}\}^m ] \end{aligned} \quad (3.3.54)$$

It follows at once from Eqs. (3.3.54) that  $p_{m,0}$  depends only on  $P$  and not on the detailed state of polarization (in contrast to  $p_{m,k}$ ). This result indicates the possibility of determining the degree of polarization from different types of experiments than are usually employed. If the field is completely polarized, the expression (3.3.54) reduces to  $p_{m,k} = \Gamma(m)$ , that is, independent of  $k$ . For unpolarized light, one gets  $p_{m,0} = \Gamma(m)/2^{m-1}$ .

### 3.3.3.3. Probability Density Functions of Normalized Stokes Parameters

We pursue further the analysis in considering the statistics of the normalized Stokes parameters. Here we are concerned with the temporal statistics of the normalized Stokes parameters for a Gaussian optical wavefield [48]. Before proceeding, we would like to emphasize several striking differences between the description of polarization in terms of normalized Stokes parameters and nonnormalized Stokes parameters.

Studies of propagation of waves in random and anisotropic media focused attention on the importance of investigating the fluctuations of the normalized instantaneous Stokes parameters  $s_j \equiv S_j/S_0$ . From Eqs. (3.2.16), it is easy to prove that

$$\sum_{j=1}^3 s_j^2 = 1 \quad (3.3.55)$$

or taking averages of both sides of Eq. (3.3.55)

$$\sum_{j=1}^3 \langle s_j^2 \rangle = 1 \quad (3.3.56)$$

which expresses the fact that the variances of  $s_j$  are linearly interrelated. We have seen that the simple topology of the Poincaré sphere can still be used to visualize states of polarization.

We now pass to the consideration of the statistics of  $s_1$ . The statistics of  $s_2$  and  $s_3$  follow from the same method by using a symmetry argument already mentioned. Now on inserting from Eqs. (3.2.13) and (3.2.14) into  $s_1 = S_1/S_0$ ,

we find that

$$s_1 = \frac{1 - (A_2/A_1)}{1 + (A_2/A_1)} \quad (3.3.57)$$

where we have set  $A_1 \equiv |E_1|^2$  and  $A_2 \equiv |E_2|^2$ . Under the conditions stated above, the PDF of  $s_1$  may be evaluated in closed form. From Eq. (3.3.5), we write the joint PDF of  $A_1$  and  $A_2$ :

$$p_{A_1, A_2}(A_1, A_2) = \frac{4}{\langle S_0 \rangle^2 (1 - P^2)} \exp\left(-\frac{2(A_1 + A_2) + \langle \sigma_1 \rangle (A_2 - A_1)}{\langle S_0 \rangle (1 - P^2)}\right) \\ \times I_0\left(\frac{4(P^2 - \langle \sigma_1 \rangle^2)^{1/2}}{\langle S_0 \rangle (1 - P^2)} (A_1 A_2)^{1/2}\right) \quad (3.3.58)$$

where  $I_0$  denotes the modified Bessel function of the first kind of order zero. Then, if we set  $X \equiv A_2/A_1$  as an auxiliary random variable, we have

$$p_X(u) = \int_{R^+} A_2 p_{A_1, A_2}(uA_2, A_2) dA_2 \quad (3.3.59)$$

On substituting from Eq. (3.3.58) into Eq. (3.3.59) and if we employ the known integral [49]

$$\int_0^\infty t \exp(-pt) I_0(at) dt = -\frac{\partial}{\partial p} \left( \int_0^\infty \exp(-pt) I_0(at) dt \right) = \frac{p}{(p^2 - a^2)^{3/2}} \quad (3.3.60)$$

the expression of  $p_X$  reduces to

$$p_X(x) = \frac{(1 - P^2)[x + 1 + \langle \sigma_1 \rangle (x - 1)]}{[(x + 1 + \langle \sigma_1 \rangle (x - 1))^2 - 4x(P^2 - \langle \sigma_1 \rangle^2)]^{3/2}} \quad (3.3.61)$$

Now the PDF of  $Z \equiv (1 - X)/(1 + X) = f(X)$  can be expressed by the formula

$$p_Z(z) = \frac{p_X(f^{-1}(z))}{|dz/dx|} \quad (3.3.62)$$

Using the definition (3.3.57) of  $s_1$  we obtain for  $p_{s_1}(s_1)$  the expression

$$p_{s_1}(s_1) = \frac{2}{(1 + s_1)^2} p_X \left( \frac{1 - s_1}{1 + s_1} \right) \quad (3.3.63)$$

Finally substitution of Eq. (3.3.61) into Eq. (3.3.63) yields the PDF of  $s_1$ :

$$p_{s_1}(s_1) = \begin{cases} \frac{(1-P^2)}{2} \left( \frac{1-s_1\langle\sigma_1\rangle}{[1-s_1\langle\sigma_1\rangle^2 - (1-s_1^2)(P^2 - \langle\sigma_1\rangle^2)]^{3/2}} \right) & \text{for } |s_1| \leq 1 \\ 0 & \text{elsewhere} \end{cases}$$

$$= \begin{cases} \frac{(1-P^2)}{2} \left( \frac{1-s_1\langle\sigma\rangle}{[1+\langle\sigma_1\rangle^2 - P^2 - 2s_1\langle\sigma_1\rangle + s_1^2P^2]^{3/2}} \right) & \text{for } |s_1| \leq 1 \\ 0 & \text{elsewhere} \end{cases} \quad (3.3.64)$$

Note that the PDF  $p_{s_2}(s_2)$  and  $p_{s_3}(s_3)$  may be evaluated from the expression of  $p_{s_1}(s_1)$  by making use of the symmetry transformations [Eqs. (3.3.46a,b)]. For instance, the corresponding PDF of  $s_2$  will read as Eq. (3.3.63) with the changes  $s_1 \rightarrow s_2$  and  $\langle\sigma_1\rangle \rightarrow \langle\sigma_2\rangle$ .

Equation (3.3.64) is sufficiently interesting to deserve some comments. First, it is important to appreciate that  $\lim_{P \rightarrow 1} p_{s_1}(s_1) = \delta(s_1 - \langle\sigma_1\rangle)$ , where  $\delta$  is the Dirac delta function. This is consistent with the fact that a pure state of polarization is a localized state in the Poincaré space, that is, characterized by a variance  $\langle s_j^2 \rangle - \langle s_j \rangle^2 = 0$ . At the other extreme of a completely unpolarized light ( $P = 0$ ) we find that  $p_{s_1}(s_1) = \frac{1}{2}$ , indicating that  $s_1$  is uniformly distributed in the basic interval  $[-1, 1]$ . Intuitively, this is what we would expect. We call such a state a *delocalized state* in conformity with the usual interpretation of the Poincaré space.

Figures 3.3.7–3.3.9 show the behavior of the PDF of the normalized Stokes parameters as a function of  $P$  for  $\langle\sigma_1\rangle = \pm \frac{1}{2}$  and  $\langle\sigma_1\rangle = 0$ . Note that for  $\langle\sigma_1\rangle = 0$ , the PDF of  $s_1$  is symmetric around its mean.

The first two moments of  $s_1$ ,

$$\langle s_1 \rangle = \int_{-1}^1 s_1 p_{s_1}(s_1) ds_1 \quad \text{and} \quad \langle s_1^2 \rangle = \int_{-1}^1 s_1^2 p_{s_1}(s_1) ds_1$$

are now computed analytically and evaluated numerically for  $\langle\sigma_1\rangle = \pm \frac{1}{2}$  and  $\langle\sigma_1\rangle = 0$ . From the definition of these moments and making use of Eq. (3.3.64), we can write

$$\langle s_1 \rangle = \frac{(1-P^2)}{2} \int_{-1}^1 \frac{x(1-x\langle\sigma_1\rangle)}{\Omega^{3/2}} dx = \frac{(1-P^2)}{2} (I_1 - \langle\sigma_1\rangle I_2) \quad (3.3.65)$$

and

$$\langle s_1^2 \rangle = \frac{(1-P^2)}{2} (I_2 - \langle\sigma_1\rangle I_3) \quad (3.3.66)$$

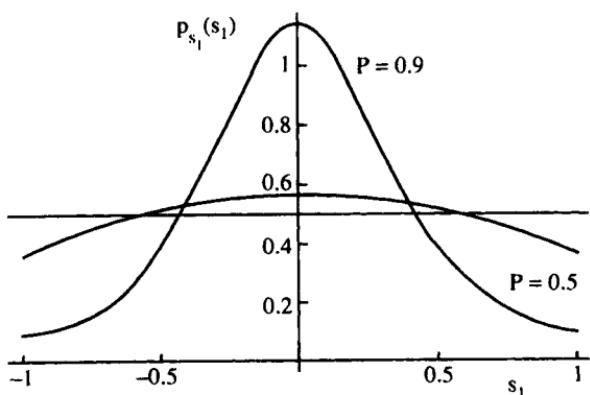


Figure 3.3.7. PDF of  $s_1$  as function of  $s_1$  for  $\langle \sigma_1 \rangle = 0$  and for different values of  $P$ .

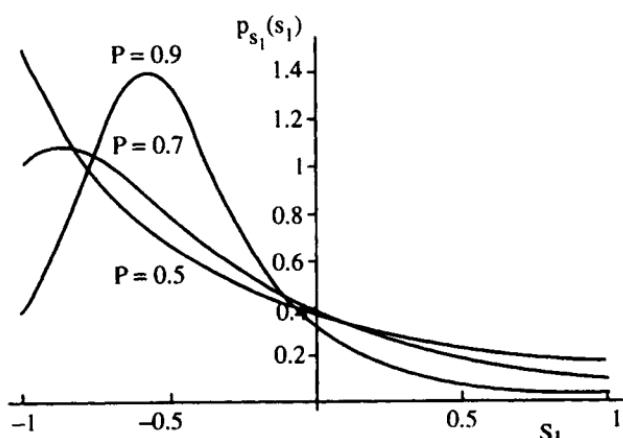


Figure 3.3.8. PDF of  $s_1$  as function of  $s_1 P$  for  $\langle \sigma_1 \rangle = -\frac{1}{2}$  and for different values of  $P$ .

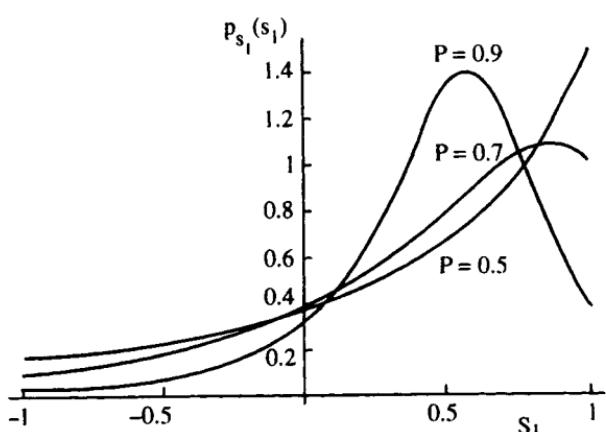


Figure 3.3.9. PDF of  $s_1$  as function of  $s_1$  for  $\langle \sigma_1 \rangle = \frac{1}{2}$  and for different values of  $P$ .

where we have set

$$I_n = \int_{-1}^1 \frac{x^n}{\Omega^{3/2}} dx \quad \text{and} \quad \Omega \equiv 1 + \langle \sigma_1 \rangle^2 - P^2 - 2x\langle \sigma_1 \rangle + x^2P^2$$

Evaluation of  $I_n$  for  $n = 1, 2, 3$  gives

$$I_1 = \frac{2\langle \sigma_1 \rangle}{(1 - P^2)(1 - \langle \sigma_1 \rangle^2)} \quad (3.3.67a)$$

$$I_2 = \frac{2(1 - P^2 - \langle \sigma_1 \rangle^2)}{P^2(1 - P^2)(1 - \langle \sigma_1 \rangle^2)} + \frac{\Delta}{P^3} \quad (3.3.67b)$$

$$I_3 = \frac{2\langle \sigma_1 \rangle(3\langle \sigma_1 \rangle^2 - 2\langle \sigma_1 \rangle^2P^2 + P^4 + 2P^2 - 3)}{P^4(1 - P^2)(1 - \langle \sigma_1 \rangle^2)} + \frac{6\langle \sigma_1 \rangle\Delta}{2P^5} \quad (3.3.67c)$$

with

$$\begin{aligned} \Delta \equiv & \sinh^{-1} \left( \frac{P^2 + \langle \sigma_1 \rangle}{\sqrt{(1 - P^2)(P^2 - \langle \sigma_1 \rangle^2)}} \right) \\ & + \sinh^{-1} \left( \frac{P^2 - \langle \sigma_1 \rangle}{\sqrt{(1 - P^2)(P^2 - \langle \sigma_1 \rangle^2)}} \right) \end{aligned} \quad (3.3.68)$$

By substituting Eqs. (3.3.67a–c) into Eqs. (3.3.65) and (3.3.66), we arrive at the expressions

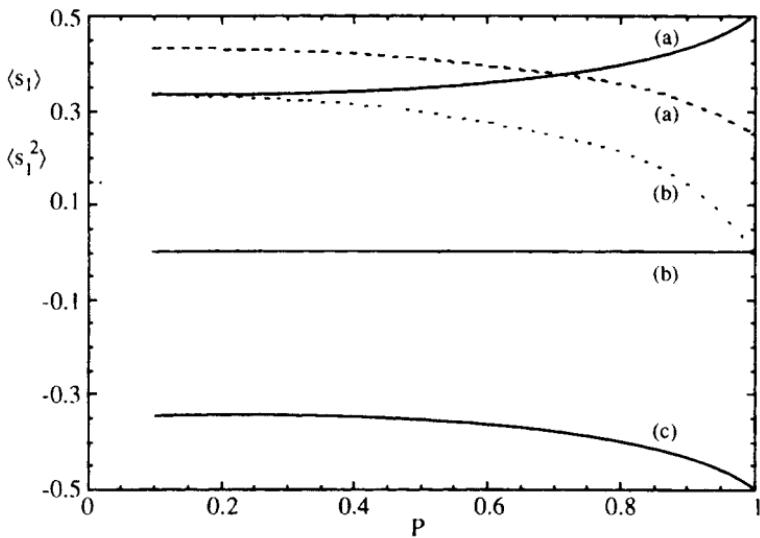
$$\langle s_1 \rangle = \frac{\langle \sigma_1 \rangle}{P^2} - \frac{\langle \sigma_1 \rangle(1 - P^2)}{2P^3}\Delta \quad (3.3.69)$$

and

$$\langle s_1^2 \rangle = \frac{\langle \sigma_1 \rangle^2}{P^4} - \frac{(P^2 - 2\langle \sigma_1 \rangle^2)(1 - P^2)}{P^4} + \frac{(P^2 - 3\langle \sigma_1 \rangle^2)(1 - P^2)}{2P^5}\Delta \quad (3.3.70)$$

Figure 3.3.10 shows the dependence of  $\langle s_1 \rangle$  [Eq. (3.3.79)] and  $\langle s_1^2 \rangle$  [Eq. (3.3.70)] as a function of the degree of polarization. Increasing the degree of polarization has the effect of lowering  $\langle s_1^2 \rangle$ . This is why the variance of  $s_j$  is minimum at  $P = 1$ .

Two appropriate remarks are appropriate at this point. First we note that an examination of Figs. 3.1.3 leads to the conclusion that the statistics of the normalized instantaneous Stokes parameters  $s_1$  is markedly different from the statistics of the instantaneous Stokes parameters  $S_1$ . Second, Eq. (3.3.64) suggests the possibility of determining the degree of polarization from different



**Figure 3.3.10.** First-order  $\langle s_1 \rangle$  and second-order moment  $\langle s_1^2 \rangle$  as functions of the degree of polarization  $P$ , for different values of  $\langle \sigma_1 \rangle$ . Solid (resp. dashed) line represents the first-order (resp. second-order) moment. Symbols: (a)  $\langle \sigma_1 \rangle = \frac{1}{2}$ , (b)  $\langle \sigma_1 \rangle = 0$ , (c)  $\langle \sigma_1 \rangle = -\frac{1}{2}$ .

types of experiments that are usually employed. We defer until Section 4.4 some illustrations of how the statistics of normalized Stokes parameters may have applications for the characterization of a multiply scattered wave by a spatially random medium. For the sake of completeness, we indicate that the conditional PDF of the normalized Stokes parameters (i.e., PDFs for the condition that a Stokes parameters is no lower than a threshold value) has been computed in Ref 50.

### 3.3.3.4. Statistics of Time-Integrated Stokes Parameters

What we want to characterize next are the statistical properties of the time-integrated Stokes parameters

$$\Omega_j(t) \equiv \int_0^t S_j(u) du$$

which are, of course, random processes. The basic quantities of interest are now the probability density functions  $p(\Omega_k)$  of the Stokes parameters integrated over a fixed time interval. No simple analytic formula is known, but what we want is an asymptotic expression for two cases of interest, namely, the cases of an integration time very long and very short compared to the coherence time  $\tau_2$  of the wavefield. We limit our discussion to the case of  $S_0$ .

When the measurement time  $t$  is short compared with  $\tau_2$ , the total instantaneous intensity of the wavefield  $S_0$  may be taken as nearly constant over the time  $t$  so that  $\Omega_0 \cong S_0 t$ . Thus  $p(\Omega_0)$  reduces to Eq. (3.3.41), that is, the PDF of the instantaneous intensity.

For large measurement times ( $t \gg \tau_2$ ), we can divide this time interval into a large number of subintervals, each of which is greater than or of the order of  $\tau_2$ . The contributions to  $\Omega_0$  for each of these subintervals are statistically independent random processes. By virtue of the central-limit theorem, it can be shown that  $\Omega_0$  is approximately Gaussian-distributed:

$$p(\Omega_0) \cong \frac{1}{(2\pi \text{var}(\Omega_0))^{1/2}} \exp\left(-\frac{(\Omega_0 - \langle\Omega_0\rangle)^2}{2\text{var}(\Omega_0)}\right) \quad (3.3.71)$$

where

$$\langle\Omega_0\rangle = \langle S_0 \rangle t \quad \text{and} \quad \text{var}(\Omega_0) = \text{var}(S_0)t^2$$

Note that

$$p(\Omega_0) \rightarrow \delta(\Omega_0 - \langle\Omega_0\rangle)$$

for extremely large values of  $t$ .

Generally the expression of  $p(\Omega_0)$  cannot be calculated explicitly and one must resort to numerical techniques. However, there is at least one spectrum for which there is a closed-form solution: the narrow Lorentzian spectral profile. In this special case and assuming very large  $t$ , Jaiswal and Mehta [51] have shown that  $p(\Omega_0)$  is given by

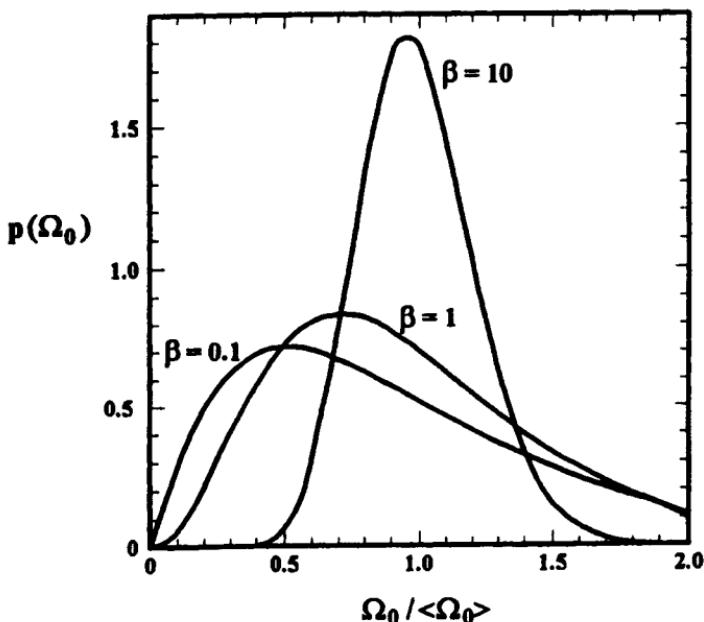
$$p(\Omega_0) = \frac{4 \exp(\beta)}{\langle\Omega_0\rangle} \sum_{m=0}^{\infty} \frac{(-1)^m \omega_m^2}{\omega_m^2 + 2\beta + \beta^2} \times \left[ \exp\left(-\frac{(\omega_m^2 + \beta^2)\Omega_0}{\beta\langle\Omega_0\rangle(1+P)}\right) \frac{Q(x_m)}{1+P} + \exp\left(-\frac{(\omega_m^2 + \beta^2)\Omega_0}{\beta\langle\Omega_0\rangle(1-P)}\right) \frac{Q(x'_m)}{1-P} \right] \quad (3.3.72)$$

where we have defined

$$Q(x) \equiv \exp(\beta) \left( \cosh(x) + \frac{1}{2} \sinh(x) \left( \frac{x}{\beta} + \frac{\beta}{x} \right) \right)^{-1},$$

$$x_m = \left( \beta^2 - \frac{1-P}{1+P} (\omega_m^2 + \beta^2) \right)^{1/2}, \quad x'_m = \left( \beta^2 - \frac{1+P}{1-P} (\omega_m^2 + \beta^2) \right)^{1/2}$$

In Eq. (3.3.72),  $\beta = 2\pi\Delta v t$  with  $\Delta v$  denoting the half-width of the Lorentz



**Figure 3.3.11.** The probability density function of the integrated intensity for a spectral Lorentz line as a function of  $\Omega_0/\langle\Omega_0\rangle$ ,  $P = 0$  for  $\beta = 0.1, 1$  and  $10$ .

spectrum. We cannot hope to give an adequate account of this involved derivation here, and must refer the reader to the paper of Jaiswal and Mehta for the full details [51]. With  $p(\Omega_0)$  in closed form an illustration of numerical results is not without interest. Figure 3.3.11 shows the behavior of  $p(\Omega_0)$  as a function of  $\Omega_0/\langle\Omega_0\rangle$  for several values of  $\beta$  and  $P = 0$ .

### 3.3.4. UNPOLARIZED RADIATION: STOKES–VERDET–BARAKAT CONDITIONS

In the preceding sections we were interested primarily in partial polarization. Now it is useful to sharpen the definition of the special case of unpolarized quasimonochromatic radiation. Most light encountered in nature is unpolarized, hence the term “natural” or “common” light, which is frequently used in the literature. In the following paragraphs we want to make more precise the sense in which quasimonochromatic light is defined as unpolarized. The earliest attempt at a quantitative analysis of the constraints that specify the mathematical description for unpolarized light was put forward by Stokes in his classic 1852 paper [52]. The full extent of the Stokes formalism for analyzing unpolarized light was pointed out in 1869, by Verdet [53]. It was mentioned previously that unpolarized light may be obtained by incoherently superposing two orthogonal pure states of polarization in equal proportions.

A second parametrization identify an unpolarized plane wave by the diagonal density matrix  $\mathbf{D}_{2u} = \frac{1}{2}\sigma_0$ . The mere fact that this density matrix is expressible in this form implies that the unpolarized state of polarization is represented by the center of the unit ball  $\Sigma_1^3$ . Here, we seek to deal with this problem by introducing necessary and sufficient conditions for light to be unpolarized. In view of our discussion in Section 3.1.1, we have anticipated that during times that are much smaller than  $\tau_2$ , natural light may be considered as completely polarized and described by an ellipse whose size, eccentricity, and orientation are randomly changed in time. Much of this section follows from Barakat's treatment of the problem, in which a more systematic and precisely stated statistical characterization of unpolarized light was derived [54].

With these comments as motivation, we begin by stating the Stokes–Verdet–Barakat conditions, namely, the necessary and sufficient conditions to guarantee that light be unpolarized. The PDF of the electric field should be (1) invariant with respect to the rotation of the axes of the reference coordinate system (i.e. symmetry about the direction of propagation), (2) invariant with respect to any phase change introduced in one of the components, and (3) independent of time.

We start by making a rotational transformation of the electric field components  $E_j(t)$  to a new coordinate frame, making an angle  $\Delta$  with the original one, and subjecting the second component to a constant phase retardation  $\delta$ :

$$E_1(t) = a_1(t) \exp(-2i\pi\nu_0 t + i\theta_1(t)) \quad (3.3.73a)$$

$$E_2(t) = a_2(t) \exp(-2i\pi\nu_0 t + i\theta_2(t) - i\delta) \quad (3.3.73b)$$

Then we have the electric field given by

$$\begin{bmatrix} E'_1(t) \\ E'_2(t) \end{bmatrix} = \begin{bmatrix} \cos(\Delta) & \sin(\Delta) \\ -\sin(\Delta) & \cos(\Delta) \end{bmatrix} \begin{bmatrix} E_1(t) \\ E_2(t) \end{bmatrix} \quad (3.3.74)$$

In the situation envisaged, the corresponding intensity projections are

$$\begin{aligned} I'_1(t) \equiv |E'_1(t)|^2 &= a_1^2(t) \cos^2(\Delta) + a_2^2(t) \sin^2(\Delta) \\ &\quad + 2a_1(t)a_2(t) \cos(\theta_1(t) - \delta) \cos(\Delta) \sin(\Delta) \end{aligned} \quad (3.3.75a)$$

$$\begin{aligned} I'_2(t) \equiv |E'_2(t)|^2 &= a_1^2(t) \sin^2(\Delta) + a_2^2(t) \cos^2(\Delta) \\ &\quad - 2a_1(t)a_2(t) \cos(\theta_1(t) - \delta) \cos(\Delta) \sin(\Delta) \end{aligned} \quad (3.3.75b)$$

where  $\theta(t) \equiv \theta_2(t) - \theta_1(t)$ . The total instantaneous intensity of the wave is given by

$$I(t) = I'_1(t) + I'_2(t) = a_1^2(t) + a_2^2(t) \quad (3.3.76)$$

Time averaging Eqs. (3.3.76) yields

$$\begin{aligned}\langle I'_1 \rangle &= \langle a_1^2 \rangle \cos^2(\Delta) + \langle a_2^2 \rangle \sin^2(\Delta) + 2(\langle a_1 a_2 \cos(\theta) \rangle \cos(\delta) \\ &\quad + \langle a_1 a_2 \sin(\theta) \rangle \sin(\delta)) \cos(\Delta) \sin(\Delta)\end{aligned}\quad (3.3.77a)$$

$$\begin{aligned}\langle I'_2 \rangle &= \langle a_1^2 \rangle \sin^2(\Delta) + \langle a_2^2 \rangle \cos^2(\Delta) - 2(\langle a_1 a_2 \cos(\theta) \rangle \cos(\delta) \\ &\quad + \langle a_1 a_2 \sin(\theta) \rangle \sin(\delta)) \cos(\Delta) \sin(\Delta)\end{aligned}\quad (3.3.77b)$$

The first moments will be independent of  $\Delta$  and  $\gamma$  if they satisfy the following conditions:

$$\begin{aligned}\langle a_1^2 \rangle &= \langle a_2^2 \rangle = \langle a^2 \rangle \\ \langle a_1 a_2 \cos(\theta) \rangle &= \langle a_1 a_2 \sin(\theta) \rangle = 0\end{aligned}\quad (3.3.78)$$

Equations (3.3.78) assert that the orthogonal field components and the phase of unpolarized light are uncorrelated. We may use Eq. (3.3.78) to rewrite Eqs. (3.3.77a,b):

$$\langle I'_1 \rangle = \langle I'_2 \rangle = \langle a^2 \rangle \quad (3.3.79)$$

Note from Eqs. (3.1.112a–d) that Eqs. (3.3.78) are the time average of the instantaneous Stokes parameters and can be rewritten as

$$\begin{aligned}\langle S_0 \rangle &= 2\langle a^2 \rangle \\ \langle S_1 \rangle &= \langle S_2 \rangle = \langle S_3 \rangle = 0\end{aligned}\quad (3.3.80)$$

in accordance with the fact that the density matrix  $\mathbf{D}_2$  of unpolarized light is a multiple of the unit matrix in that case.

Let us now examine the higher moments of the two projected intensities

$$\begin{aligned}\langle (I'_1)^2 \rangle &= 2\langle a_1^2 \rangle^2 \cos^4(\Delta) + 2\langle a_2^2 \rangle^2 \sin^4(\Delta) + 2\langle a_1^2 a_2^2 \rangle \cos^2(\Delta) \sin^2(\Delta) \\ &\quad + 4[\langle a_1^3 a_2 \cos(\theta) \rangle \cos(\delta) + \langle a_1^3 a_2 \sin(\theta) \rangle \sin(\delta)] \cos^3(\Delta) \sin(\Delta) \\ &\quad + 4[\langle a_1 a_2^3 \cos(\theta) \rangle \cos(\delta) + \langle a_1 a_2^3 \sin(\theta) \rangle \sin(\delta)] \cos(\Delta) \sin^3(\Delta) \\ &\quad + 4[\langle (a_1 a_2 \cos(\theta))^2 \rangle \cos^2(\delta) + \langle (a_1 a_2 \sin(\theta))^2 \rangle \sin^2(\delta) \\ &\quad + 2\langle a_1^2 a_2^2 \cos(\theta) \sin(\theta) \rangle \cos(\delta) \sin(\delta)] \cos^2(\Delta) \sin^2(\Delta)\end{aligned}\quad (3.3.81)$$

with a similar formula for  $\langle (I'_2)^2 \rangle$ . The right-hand side of Eq. (3.3.81) can be simplified further by using Eqs. (3.1.112a–d) and (3.3.7). We have

$$\langle (I'_1)^2 \rangle = 2\langle a^2 \rangle^2 \quad (3.3.82)$$

Evaluating  $\langle(I'_2)^2\rangle$  leads to the same result; hence

$$\langle(I'_j)^2\rangle = \frac{1}{2}\langle I' \rangle^2 = \frac{1}{2}\langle S_0 \rangle^2 \quad (3.3.83)$$

Under these conditions Barakat [54] has shown, using a symbol manipulation package, that

$$\langle(I'_j)^l\rangle = \Gamma(l + 1)\langle a^2 \rangle^l \quad (3.3.84)$$

Consequently all moments and products moments of the two projected intensities  $I'_1$  and  $I'_2$  are independent of the angle of rotation  $\Delta$  of the axes of the reference coordinate system, of the phase retardation  $\gamma$  introduced into one of the components, and of the time. These moments characterize a negative exponential PDF:

$$p(I'_j) = \frac{1}{\langle a^2 \rangle} \exp\left(-\frac{I'_j}{\langle a^2 \rangle}\right) \quad (3.3.85)$$

It can be shown (see details in Barakat [54]) that  $I'_1$  and  $I'_2$  are statistically independent:

$$\langle(I'_1)^k(I'_2)^l\rangle = \langle(I'_1)^k\rangle\langle(I'_2)^l\rangle = \Gamma(l + 1)\Gamma(k + 1)\langle a^2 \rangle^{k+l} \quad (3.3.86)$$

or equivalently

$$p(I'_1, I'_2) = p(I'_1)p(I'_2) \quad (3.3.87)$$

The expressions in Eqs. (3.3.84)–(3.3.87) constitute the Stokes–Verdet–Barakat theorem [54]. In our opinion, this theorem tells one relatively little about the underlying physics of unpolarized light because it does not reflect the evolutiveness of the instantaneous state. We therefore conclude this section with some comments regarding an alternative approach to the modeling of unpolarized quasimonochromatic light via a random walk. A step was recently taken in this direction [55]. Potekhin relates radiation fluctuations to properties of random walkers on the Poincaré sphere  $\Sigma_1^2$ . These fluctuations lead to a fuzziness in the position of the tip of the Stokes vector. Thus the behavior of the system can be described by a fuzzy trajectory on the Poincaré sphere  $\Sigma_1^2$  with the size of the fuzz indicating the extent of the fluctuations, specifically, the partially polarized character of the wave. Unpolarized light corresponds to a random walk of the position of the tip uniformly distributed over  $\Sigma_1^2$ . This random-walk approach complements the results of the statistics of the normalized instantaneous Stokes parameters (Section 3.3.3).

### 3.3.5. POLYCHROMATIC RADIATION WAVEFIELD

The developments introduced in the preceding sections have all concerned light of arbitrary narrow spectral extent. In some cases of practical interest, the simplifying assumption of quasimonochromatic light is no longer valid, that is,  $\Delta\nu$  is not extremely small compared with the mean frequency  $\nu_0$ , and we assume that the wave is of arbitrary spectral distribution. For such a case Pancharatnam has shown that the elements of the coherency matrix for a polychromatic radiation field are simply the sum of the corresponding elements for all frequency components [56].

According to Eqs. (3.1.9) and (3.1.75), it follows that we may write for the coherency matrix

$$\Phi_{jk}(\mathbf{R}, \mathbf{R}, 0) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \int_0^\infty \int_0^\infty \hat{E}_j(\mathbf{R}, \nu_1; T) \hat{E}_k^*(\mathbf{R}, \nu_2; T) \\ \times \exp(-2i\pi\nu_1 t) \exp(2i\pi\nu_2 t) d\nu_1 d\nu_2 dt \quad (3.3.88)$$

Now, interchanging the order of integration and time integrating yields

$$\Phi_{jk}(\mathbf{R}, \mathbf{R}, 0) = \lim_{T \rightarrow \infty} \int_0^\infty \int_0^\infty \hat{E}_j(\mathbf{R}, \nu_1; T) \hat{E}_k^*(\mathbf{R}, \nu_2; T) \operatorname{sinc}(2\pi(\nu_2 - \nu_1)T) d\nu_1 d\nu_2 \quad (3.3.89)$$

Next we make use of the expression

$$\lim_{T \rightarrow \infty} \operatorname{sinc}(2\pi(\nu_2 - \nu_1)T) = \delta(\nu_2 - \nu_1) \quad (3.3.90)$$

which is one of the many approximations to the delta function. We finally obtain the expression for the coherency matrix of the light:

$$\Phi_{jk}(\mathbf{R}, \mathbf{R}, 0) = \int_0^\infty \hat{\Phi}_{jk}(\mathbf{R}, \mathbf{R}, \nu) d\nu \quad (3.3.91)$$

Hence the proof is completed. Equation (3.3.91) is the central result of the analysis and reflects the fact that the different frequency components are uncorrelated with respect to one another. For fields of arbitrary spectral distribution, the polarization and spatial coherence must be regarded as spectrally distributed quantities.

## SECTION 3.4

# Entropy of the Radiation Field

The entropy of a radiation field is one of the canonical problems of statistical physics and has attracted much attention in the past. To our knowledge, the subject has a long history that has been traced by Wien [57] in a classic paper, published as far back as 1894, and it has culminated with seminal papers by Planck [58] and von Laue [59] offering a number of exciting insights into the development of the fundamental quantization of energy. Since these pioneering works, important advances were discussed by Clark Jones [60] and O'Neill [61]. Some further commentary on the historical background to this development is given in Appendix E. However, very little of the insight of these pioneering works has been brought to bear on the problem of polarization, and their implications are not always understood in a quantitative fashion. Therefore, the need for a more consistent theory as led us to begin the development of a thermodynamic approach to the problems of polarization optics.

The primary purpose of this chapter is to provide a detailed analysis of those theoretical aspects of the theory of radiation field able to account for the entropy of a stochastic partially polarized field. We shall try to give the term a precise meaning when applied to electromagnetic fields. The seminal concept underlying the work reported here is the concept of scalar invariant, that is, the generalization of the degree of polarization. This parameter has a statistical as well as geometric significance. The former follows from the second-order description of the polarization density matrix. The latter, which may be visualized within the Poincaré space, is of the nature of an order parameter and arises as a direct consequence of the algebraic structure of the density matrix representation for the underlying  $SU(N)$  group. The secondary purpose is to examine some of the broad general features of radiation entropy from a thermodynamical point of view.

This section is organized as follows. In Section 3.4.1 we begin by addressing the important question of precisely what the entropy of polarization actually means operationally and, in so doing, we endeavor to explain physically why the entropy is under the dependence of a set of scalar invariants; in particular, we find that a number of important features can be found as a direct consequence of the  $SU(N)$  Lie group expansion of the polarization density matrix. We shall derive a closed expression for the polarization dependence of radiation entropy. A thermodynamical analysis underlying the two-level

description of a partially polarized wave is presented in Section 3.4.2. In the thermodynamic description the temperature plays a prominent role. With this in mind, we shall define an effective polarization temperature. Finally, in Section 3.4.3 we will discuss, on the basis of the maximum entropy principle, the problems associated with extending the treatment of a plane wave to a non-plane-wave radiation field. Entropy considerations play a prominent role in the analysis of the irreversible evolution of multiply scattered light by a spatially random medium.

### 3.4.1. ENTROPY OF A STOCHASTIC PLANE WAVEFIELD

We consider a quasimonochromatic optical field in the form of plane waves, of infinite extent, propagating in some direction characterized by the unit vector  $\mathbf{e}_3$ . The transverse field is resolved into two orthogonal components along the directions characterized by the unit vectors  $\mathbf{e}_1$  and  $\mathbf{e}_2$  corresponding to orthogonal linear polarizations. All matrix quantities will be defined in this linear polarization basis.

Our starting point is the von Neumann entropy  $S$  of the radiation field in the impure (mixed) state represented by the density matrix  $\mathbf{D}_2$ . It is defined according to the usual dimensionless version appearing in statistical mechanics as

$$S = -\text{tr}(\mathbf{D}_2 \ln(\mathbf{D}_2)) \quad (3.4.1)$$

where the density matrix is as given by Eq. [62, 63]. The von Neumann entropy is a quantitative measure of the amount of information that would be gained by switching from a mixed state to a pure state. This orthogonal decomposition is the most economical representation of  $\mathbf{D}_2$  (in the sense of entropy minimization) [64]. Several properties of  $S$  are of importance for us. First, it is worth observing that Eq. (3.4.1) is basis-independent; the entropy remains invariant under a similarity transformation of the density matrix  $\mathbf{D}_2 \rightarrow \mathbf{R}\mathbf{D}_2\mathbf{R}^{-1}$ . This is expected since polarization properties must be unaffected by the particular choice of basis. Second, we note that the mapping  $\mathbf{D}_2 \rightarrow S(\mathbf{D}_2)$  is concave; the entropy of a mixed state is greater than the constituent entropies weighted as in the mixing. Consequently, taking linear combinations  $\sum_j p_j \mathbf{D}_{2j}$  of density matrices  $\mathbf{D}_{2j}$  with real positive coefficients  $0 \leq p_j \leq 1$  summing to unity  $\sum_j p_j = 1$ , we have

$$\sum_j p_j S(\mathbf{D}_{2j}) \leq S(\mathbf{D}_2) \leq \sum_j p_j S(\mathbf{D}_{2j}) - \sum_j p_j \ln(p_j) \quad (3.4.2)$$

This equation is an optimally inequality in the sense that equality holds on the left if all  $\mathbf{D}_{2j}$  are equal, and on the right if all  $\mathbf{D}_{2j}$  have disjoint support. An interesting physical application of Eq. (3.4.2) is given by the discussion of the

entropy of an incoherent mixture of partially polarized light beams (see Appendix C). Let us remark in passing that both matrices  $\mathbf{D}_2$  and  $\ln(\mathbf{D}_2)$  are diagonalized by the same unitary transformations; this comes from the fact that  $\mathbf{D}_2$  commutes with  $\ln(\mathbf{D}_2)$ . We must note that a spectral entropy can be defined in a completely similar manner by permuting  $\mathbf{D}_2$  and  $\hat{\mathbf{D}}_2$  in Eq. (3.4.1). This is essentially the result obtained by Barakat and Brosseau, using Eq. (3.1.136b) [65].

Given these background remarks, we offer two different ways by which to obtain the degree of polarization dependence of the radiation entropy. Actually both methods complement each other in that the first approach involves an eigenvalue problem, whereas the second approach involves a geometric property of the set of polarization states. To begin with, we express Eq. (3.4.1) differently in the representation in which  $\mathbf{D}_2$  is diagonal. Denoting by  $\lambda_j$  the eigenvalues of  $\mathbf{D}_2$ , we can express the entropy, via Eq. (3.4.1), as

$$S = - \sum_j \lambda_j \ln(\lambda_j) \quad (3.4.3)$$

Our problem is reduced to the calculation of these eigenvalues. This is accomplished from the expansion of the density matrix in Eq. (3.1.133). We find

$$\lambda_j = \frac{1 \pm P}{2} \quad (3.4.4)$$

where  $P$  is the degree of polarization of the plane wave. Now by substituting Eq. (3.4.4) into Eq. (3.4.3), the result takes a relatively simple form, namely

$$S(P) \equiv -\ln(s(P)) \quad (3.4.5)$$

with

$$s(x) = \frac{1}{2}(1-x)^{(1-x)/2}(1+x)^{(1+x)/2} \quad (3.4.6)$$

This establishes the nonlinear dependence of  $P$  on  $S$ ; entropy is a measure of purity of the states of polarization [28, 65]. Equation (3.4.5) demonstrates that one needs only a single parameter,  $P$ , to characterize the entropy of a radiation field in the form of plane waves. The function  $S(P)$  is plotted graphically in Fig. 3.4.1. As  $P$  ranges from 0 to 1,  $S(P)$  goes from 0 to  $\ln(2)$  inclusive.

The curve in Fig. 3.4.1 undergoes a monotonic decrease as the degree of polarization is increased. It does not differentiate pure states; thus, irrespective of the form of the pure state of polarization considered, we have  $S(\leftrightarrow) = S(\uparrow) = S(\nearrow) = S(\nwarrow) = S(\circlearrowleft) = S(\circlearrowright) = S(P=1) = 0$ . As pointed out earlier, this property reflects the isotropy of the Poincaré sphere.

To calculate entropy, we can also take a rather different approach based on the convexity property of the states of polarization. This second proof of Eq.

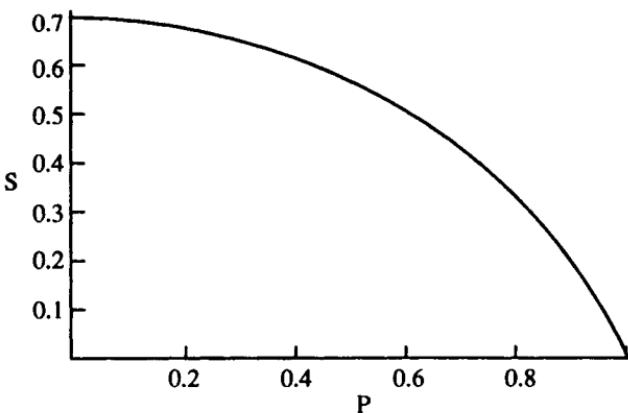


Figure 3.4.1. Dependence of the entropy  $S$  versus degree of polarization  $P$ .

(3.4.5) illuminates its significance from a geometric point of view. To do this, we start with the observation that one can always expand, uniquely, any mixed state into a purely polarized part and an unpolarized part. The relative weight of each component is determined by the degree of polarization  $P$ . Following Eq. (3.1.144), we write the density matrix of any mixed state by the convex sum

$$\mathbf{D}_2 = P\mathbf{D}_{2p} + (1 - P)\mathbf{D}_{2u} \quad (3.4.7)$$

with  $\mathbf{D}_{2u} = \frac{1}{2}\sigma_0$ . Once this is done, it becomes straightforward to derive the entropy. On substituting Eq. (3.4.7) into Eq. (3.4.1), we arrive again at Eq. (3.4.5). We also refer the interested reader to Appendix B for the details of the derivation of Eq. (3.4.5) by using the spectral decomposition theorem.

In closing this section, let us note that the entropy vanishes for a coherent wavefield that is stationary. In other words, a purely monochromatic wavefield has minimum entropy.

### 3.4.2. TEMPERATURE OF POLARIZATION

While everything appears straightforward in the procedure of obtaining the polarization dependence of the radiation entropy, we are led naturally to inquire whether a more general thermodynamic treatment of partial polarization exists. Here we detail the basic thermodynamical properties of the polarization density matrix. As an application of this approach, we indicate the relationship between the degree of polarization and the “temperature of polarization” [66].

We first wish to introduce the analogy of the two-level description of a partially polarized wave with a one-dimensional Ising spin system in contact with a heat bath and in zero magnetic field. As was discovered 60 years ago,

the Ising model is a simple model of a ferromagnet. In a ferromagnetic material, neighboring spins tend to be parallel to each other. A pair of spins has the exchange energy  $-J$ , if it is parallel, and  $J$  if it is antiparallel. The Hamiltonian assigned to a particular configuration of  $N$  spins is  $-J\sum_{\langle ij \rangle} S(\mathbf{R}_i) \cdot S(\mathbf{R}_j)$ , with each site  $\mathbf{R}_i$  having a spin  $S(\mathbf{R}_i) = \pm 1$ . The term  $\langle ij \rangle$  refers to nearest-neighbor  $i$  and  $j$  sites, and  $J$  stands for the spin-spin ferromagnetic coupling ( $J > 0$ ). A standard known result of statistical mechanics of Ising systems yields the entropy per spin [67]

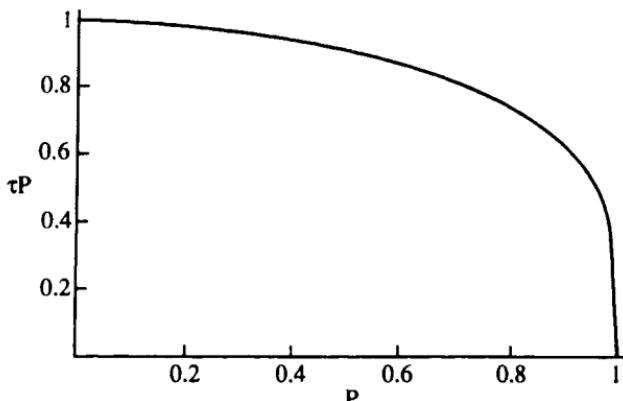
$$\frac{S(y)}{Nk} = \ln(2 \cosh(y)) - y \tanh(y) \quad (3.4.8)$$

where  $y \equiv J\beta$ ,  $\beta \equiv 1/kT$  is the inverse temperature. The constant  $k$  is the Boltzmann constant. By a straightforward calculation, Eq. (3.4.5) can be verified to be identically equal to (3.4.8) if one makes use of the following expression

$$\frac{1}{\tau} = \frac{1}{2} \left( \ln \left( \frac{1+P}{1-P} \right) \right) \quad (3.4.9)$$

and sets  $\tau \equiv kT/\Gamma$ , which defines an effective polarization temperature. A feature to be particularly emphasized in connection with Eq. (3.4.9) is that it should not be confused with the radiance temperature obtained using Planck's spectral law. Figure 3.4.2 shows  $\tau P$  as a function of  $P$ . We first observe that  $\tau$  is a monotonic decreasing function of  $P$ .

Equation (3.4.9) also shows clearly that  $\tau \sim P^{-1}$  for very small values of  $P$ , and we observe a dramatic change as  $P$  approaches 1 from below. In the limit



**Figure 3.4.2.** Polarization temperature  $\tau$  as a function of degree of polarization  $P$ . See Eq. (3.4.9).

of  $P \rightarrow 1$ , the polarization temperature may be evaluated as  $\tau \sim P^{-1}(1 - P)^\alpha$ . The exponent  $\alpha$  is found to be equal to 0.225. As one might expect from the third law of thermodynamics,  $S \rightarrow 0$  as  $\tau \rightarrow 0$ . For unpolarized light, on the other hand,  $\tau \rightarrow \infty$  implies that entropy considerations now become of paramount importance. In thermodynamics, an important role of temperature is that it allows a system to explore "phase space." For the  $SU(2)$  problem, the available phase space is, in fact, bounded on the Poincaré sphere. Such an approach is consistent with the description of unpolarized quasimonochromatic light as a random walk on the Poincaré sphere  $\Sigma_1^2$ .

At this point two comments are in order. Just as in the study of the Ising system, one can determine the thermodynamic functions for the partially polarized radiation field. Using the thermodynamics of the canonical ensemble (in the language of statistical mechanics), the partition function is given by

$$Z = 2(1 - P^2)^{-1/2} \quad (3.4.10)$$

This equation can be used to find other thermodynamic quantities such as the internal energy  $U$  and Helmholtz free energy  $F \equiv U - TS$  (see Table 3.4.1). In like fashion as one defines in thermodynamics the equilibrium temperature such as  $1/T \equiv \partial S/\partial U$ , one can see explicitly that the polarization temperature verifies  $1/\tau \equiv \partial S/\partial U = -(\partial S/\partial P)$ . In this way of looking at things, the "specific

**TABLE 3.4.1. Comparison of Thermodynamic Functions for Ising Spin- $\frac{1}{2}$  System and Partially Polarized Radiation<sup>a</sup>**

Ising Spin System	Quantity	Partially Polarized Plane Wavefield
$2 \cosh(y)$	Partition function	$2(1 - P^2)^{-1/2}$
$-J \tanh(y)$	Internal energy	$-P$
$k[\ln(2 \cosh(y))$	Entropy	$-\ln\left[\frac{1}{2}(1 + P)^{(1+P)/2}(1 - P)^{1-P/2}\right]$
$-y \tanh(y)]$		
$-\frac{y}{J} \ln[2 \cosh(y)]$	Free energy	$\frac{1}{4} \ln\left(\frac{1+P}{1-P}\right) \ln\left[\left(\frac{1+P^2}{2}\right)\right]$
$y^2[1 - \tanh^2(y)]$	Specific heat	$\frac{1}{4}(1 - P^2) \ln\left(\frac{1+P}{1-P}\right)$
$a[ \ln(\tanh(y)) ]^{-1}$	Correlation length	$l[ \ln(P) ]^{-1}$

<sup>a</sup>The term  $a$  is the lattice constant of the spin model. We have set  $\beta J \equiv y$ . Within the spin model, functions have been normalized to the number of spins. Thermodynamic functions have similar expressions provided thermodynamic temperature  $T \equiv J/k_y$  is changed into polarization temperature  $\tau \equiv \frac{1}{2} \ln[(1 + P)/(1 - P)]$ .

heat" can be found by the simple expression

$$C = \frac{1 - P^2}{\tau^2} \quad (3.4.11)$$

Specifically, it is easy to show that the "specific heat" has a maximum at a degree of polarization  $P_0$ , such that

$$P_0 \ln \left( \frac{1 + P_0}{1 - P_0} \right) = 2$$

that is,  $P_0 = \tau_0 = 0.834$ . This behavior is reminiscent of the two-level [ $SU(2)$ ] description of partial polarization. It is instructive for later purposes to push this analogy further by defining a correlation length,  $\xi$ , which indicates the range over which one spin appreciably influences the orientations of other spins. To this end we consider that the dipolar coupling falls off exponentially with spin separation

$$\langle S(\mathbf{R}_i) \cdot S(\mathbf{R}_j) \rangle \sim \exp \left( -\frac{|\mathbf{R}_i - \mathbf{R}_j|}{\xi} \right), \quad |\mathbf{R}_i - \mathbf{R}_j| \rightarrow \infty \quad (3.4.12)$$

On making the substitution, we found that the correlation length behaves as  $|\ln(P)|^{-1}$ . At small  $P$ , we therefore expect the correlations to decay to zero.

### 3.4.3. MAXIMUM ENTROPY PRINCIPLE

The maximum entropy principle (MEP) is a self-contained optimization technique that is widely used in statistical physics [64]. An attractive feature of this method of lies in its generality, and it is often viewed as a central organizing principle for statistical physics. This section is intended to serve a dual purpose. First, we wish to explain in terms as simple as possible the utility of this method in the practical calculation of the entropy of radiation. Second, the MEP provides an instructive way to show the utility of the concept of scalar invariants. Applications to plane and nonplane waves will be examined. The importance that the MEP plays in several polarization related problems such as the study of multiple scattering of light by a spatially random medium will be discussed further in Section 4, Part 4. For mathematical and historical details of the MEP, as well as various methods for implementing it, we refer the reader to Refs. 28 and 64.

The idea here is that the knowledge of a sequence of observables  $\langle \mathbf{O}_j^{(N)} \rangle$  evaluated at a particular point in space will impose a particular form of the density matrix. Let us now proceed with the mathematics of this problem. In

this section, even though we will be concerned entirely with two- and three-dimensional polarization density matrices, we will nevertheless describe the MEP for systems of arbitrary dimension that will allow us to introduce the parametrization of the system that serves as a basis of our entire treatment. Everything of interest to us for reaching an analytic expression of entropy follows from a variational principle, that is, to find the stationary values of

$$-\text{tr}(\mathbf{D}_N \ln(\mathbf{D}_N)) - \sum_{j=0}^M \text{tr}(\mathbf{D}_N \mathbf{O}_j^{(N)}) \quad (3.4.13)$$

Applying the method of Lagrange for maximizing Eq. (3.4.13) subject to the  $M (= N^2 - 1)$  data constraints, the  $N \times N$  density matrix will then be of the form

$$\mathbf{D}_N = \frac{1}{Z_N} \exp\left(-\sum_{j=1}^M \lambda_j^{(N)} \mathbf{O}_j^{(N)}\right) \quad (3.4.14)$$

where the  $M$  Lagrange multipliers  $\lambda_j^{(N)}$  (also termed the *generalized potentials*) are determined in order to satisfy  $\langle \mathbf{O}_j^{(N)} \rangle \equiv \text{tr}(\mathbf{O}_j^{(N)} \mathbf{D}_N)$  for  $j = 1, M$ . From the trace one requirement for the density matrix, one finds

$$Z_N = \text{tr}\left(\exp\left(-\sum_{j=1}^M \lambda_j^{(N)} \mathbf{O}_j^{(N)}\right)\right) \quad (3.4.15)$$

The Lagrange multipliers  $\lambda_j^{(N)}$  conjugate to the  $\mathbf{O}_j$  terms are to be chosen such

$$-\frac{1}{Z_N} \frac{\partial Z_N}{\partial \lambda_j^{(N)}} = -\frac{\partial \ln(Z_N)}{\partial \lambda_j^{(N)}} = \Theta_j^{(N)} \quad \text{for } j \in \langle 1, N \rangle \quad (3.4.16)$$

which follows from substitution of Eq. (3.4.14) into Eq. (3.1.125). The net result is the expression for the entropy, namely

$$S_N = \ln(Z_N) + \sum_{j=1}^M \lambda_j^{(N)} \Theta_j^{(N)} \quad (3.4.17)$$

Note that Eq. (3.4.17) tells us that the Lagrange multipliers reduce to

$$\lambda_j^{(N)} = \frac{\partial S_N}{\partial \Theta_j^{(N)}} \quad (3.4.18)$$

There are two important observations to be drawn from these results. Taking the determinant of both sides of Eq. (3.4.14), then making use of the formula  $\det(\exp(A)) = \exp(\text{tr}(A))$ , the linearity of the trace and Eq. (3.1.176a), we arrive

at the relation

$$Z_N = (\det(\mathbf{D}_N))^{-1/2} \quad (3.4.19)$$

This equation is sufficiently interesting to warrant some comment. Two limiting cases are of special interest. On one hand, we find for an unpolarized optical field that  $\mathbf{D}_N = (1/N)\mathbf{O}_0^{(N)}$  and  $\Theta_j^{(N)} \equiv \delta_{j0}$ . Substituting in Eq. (3.4.19), we have  $Z_N = N$ , so that finally Eq. (3.4.17) reduces to  $S_N = \ln(N)$ . On the other hand, an idempotent density matrix implies a diverging  $Z_N$ , and we have  $S_N = 0$ .

Let us examine the meaning of Eq. (3.4.17) for a few small values of  $N$ . We begin by treating the case  $N = 2$ . We consider an infinitely extended plane quasimonochromatic wave propagating along the direction characterized by the unit vector  $\mathbf{e}_3$ , which together with the directions  $\mathbf{e}_1$  and  $\mathbf{e}_2$ , form an orthogonal set. The  $\mathbf{O}_j^{(2)}$  values are the Pauli matrices and  $\Theta_j^{(2)} = \langle \sigma_j \rangle$ . In that case, the only ( $N - 1 = 1$ ) rotational invariant is

$$\text{tr}((\mathbf{D}_2)^2) = 1 - 2 \det(\mathbf{D}_2) = \frac{1 + (P_2^{(2)})^2}{2} \quad (3.4.20)$$

where  $P_2^{(2)} = (\sum_{j=1}^3 (\Theta_j^{(2)})^2)^{1/2}$  is the usual degree of polarization for plane waves. Replacing Eq. (3.4.20) in Eq. (3.4.19), we obtain for  $Z_2$

$$Z_2 = \frac{2}{(1 - (P_2^{(2)})^2)^{1/2}} \quad (3.4.21)$$

identical to the partition function Eq. (3.4.10).

By making use of Eq. (3.4.16), we find

$$\lambda_j^{(2)} = \frac{\langle \sigma_j \rangle}{2P_2^{(2)}} \ln \left( \frac{1 - P_2^{(2)}}{1 + P_2^{(2)}} \right) \quad (3.4.22)$$

If we substitute from Eq. (3.4.22) into Eq. (3.4.17), we arrive at the relation

$$S_2(P_2^{(2)}) = \ln(2) - \frac{1}{2} \ln(1 - (P_2^{(2)})^2) + \frac{1}{2} P_2^{(2)} \ln \left( \frac{1 - P_2^{(2)}}{1 + P_2^{(2)}} \right) \quad (3.4.23)$$

A quick algebraic manipulation clearly reproduces Eq. (3.4.5), specifically  $S_2(P_2^{(2)}) = -\ln(s(P_2^{(2)}))$  [28].

Figure 3.4.3 shows  $S_2$  as a function of the scalar invariant  $\text{tr}((\mathbf{D}_2)^2)$ . Minimum entropy states define the pure states, ( $P_2^{(2)} = 1$ ,  $\det(\mathbf{D}_2) = 0$ ), and are located at the surface of the so-called Poincaré sphere  $\Sigma_1^2$ . The other interesting limit is obtained by considering the maximum entropy state corresponding to the completely unpolarized state, ( $P_2^{(2)} = 0$ ,  $\mathbf{D}_2 = \sigma_0/2$ ) located at the center of the ball  $\Sigma_1^3$ . In all other cases such as  $0 < S < \ln(2)$ , mixed states are the points inside the ball  $\Sigma_1^3$ .

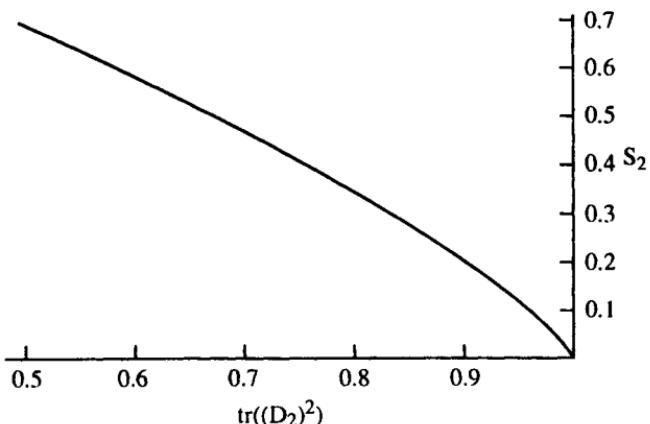


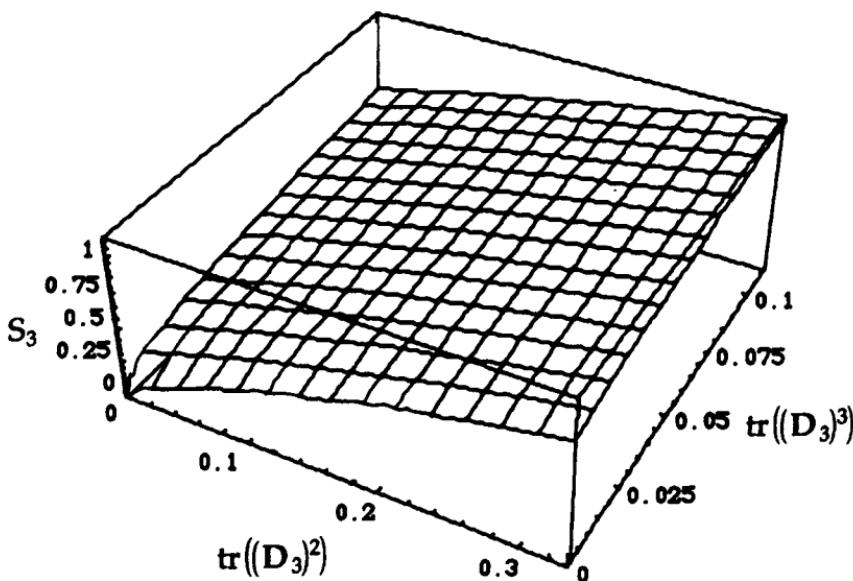
Figure 3.4.3. Plot of the entropy  $S_2$  as function of scalar invariant  $\text{tr}((\mathbf{D}_2)^2)$ .

We now pass to a consideration of the case  $N = 3$ . To this point we have treated the case of plane waves ( $N = 2$ ). This is to say that we represent the density matrix of any mixed state by a  $2 \times 2$  matrix. This description is inaccurate whenever the spatial structure of the quasimonochromatic wave is assumed arbitrary. It is this fact that makes the MEP analysis of the case  $N = 3$  particularly interesting. In contrast to plane waves that can be described, at a fixed point in space, by two transverse components of the electric field, the case of nonplanar waves requires us to consider a longitudinal component  $E_3$ .

For the case at hand, the polarization density matrix  $\mathbf{D}_3$  contains (at most) eight independent real parameters and may be determined by eight independent measurements. The eight scalar quantities contained in  $\mathbf{D}_3$  provide a complete description of an arbitrary pure or mixed state. While all 15 quantities are required to describe a mixed state, they are partially redundant for a pure state. Some reflection on the matter suggests that there now exists an analogy of the analysis of polarization for a nonplane wavefield with the statistical description of a spin-1 particle.

In view of the MEP treatment of the previous section, two order parameters are now required ( $N - 1 = 2$ ). Consequently two degrees of polarization may be defined in terms of the scalar invariants  $\text{tr}((\mathbf{D}_3)^2)$  and  $\text{tr}((\mathbf{D}_3)^3) = 3 \det(\mathbf{D}_3)$ . The entropy  $S_3$  was numerically evaluated. Figure 3.4.4 shows the three-dimensional plot of the entropy as a function of scalar invariants  $\text{tr}((\mathbf{D}_3)^2)$  and  $\text{tr}((\mathbf{D}_3)^3)$ . An interesting feature to observe in Fig. 3.4.4 is the asymmetry of this plot.

The general theory of polarization entropy expounded in this section is applicable to a variety of situations. In Part 4 of this book we shall discuss the bearing of these results on the problems of calculating the entropy transformations when light in an arbitrary state of polarization interacts with some optical element or propagates through a scattering medium [66].



**Figure 3.4.4.** Three-dimensional surface plot of the entropy  $S_3$  as function of scalar invariants  $\text{tr}((\mathbf{D}_3)^2)$  and  $\text{tr}((\mathbf{D}_3)^3)$ .

In summary, we have concentrated on describing the conceptual foundations of the theory of partial polarization of electromagnetic radiation fields. Hence, the basic definitions and the main equations were discussed in detail. The rich and original theories that have successfully evolved over the past decades are indicative of the fundamental health and future promise of the contemporary theory of partial polarization. In the next part, we show its implementation problem to determine how the state of polarization of a wave changes when it interacts with matter.

## Further Reading

1. L. Mandel and E. Wolf, *Rev. Mod. Phys.* **37**, 1965, 231. See also L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*, Cambridge Univ. Press, New York, 1995.
2. J. W. Goodman, *Statistical Optics*, Wiley, New York, 1985.
3. D. Gabor, *J. Inst. Elect. Eng.* **93**, 1946, 429. For a detailed presentation of the properties of analytic signals, we refer the reader to J. Dugundji, *IRE Trans. Info. Theory* **IR-4**, 1958, 53.
4. A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York, 1984.
5. D. Middleton, *An Introduction to Statistical Communication Theory*, McGraw-Hill, New York, 1960.
6. A. Blanc-Lapierre and R. Fortet, *Theory of Random Functions*, Gordon & Breach, New York, 1968.
7. E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, Clarendon Press, Oxford, 1948.
8. M. Beran and G. B. Parrent, Jr., *Theory of Partial Coherence*, Prentice-Hall, Englewood Cliffs, NJ, 1964.
9. R. M. A. Azzam and N. B. Bashara, *Ellipsometry and Polarized Light*, North-Holland, Amsterdam, 1977. See also D. A. Holmes and D. L. Feucht, *J. Opt. Soc. Am.* **57**, 1967, 466.
10. A. S. Marathay, *Elements of Optical Coherence Theory*, Wiley, New York, 1982.
11. N. Wiener, *Acta Math.* **55**, 1930, 117. See also N. Wiener, *J. Math. Phys. (MIT)* **2**, 1923, 131.
12. A. Y. Khintchine, *Math. Ann.* **109**, 1934, 605.
13. R. C. Bourret, *Nuovo Cimento* **18**, 1960, 347.
14. Y. Kano and E. Wolf, *Proc. Phys. Soc.* **80**, 1962, 1273.
15. C. L. Mehta and E. Wolf, *Phys. Rev.* **A134**, 1964, 1143.
16. R. Barakat, *J. Opt. Soc. Am.* **A4**, 1987, 1256. See also R. Barakat, *Optica Acta* **32**, 1985, 295.
17. R. Barakat, *J. Opt. Soc. Am.* **53**, 1963, 317.
18. E. Wolf, *Nuovo Cimento* **13**, 1959, 1165. See also E. Wolf, *Nuovo Cimento* **12**, 1954, 884.
19. G. B. Parrent, Jr., *J. Opt. Soc. Am.* **49**, 1959, 787.

20. P. Roman, *Nuovo Cimento* **13**, 1959, 974. See also P. Roman and E. Wolf, *Nuovo Cimento* **17**, 1960, 462.
21. For an account of the density matrix theory, see K. Blum, *Density Matrix Theory and Applications*, Plenum Press, New York, 1981.
22. C. Brosseau and R. Barakat, *Opt. Commun.* **91**, 1992, 408.
23. The concept of cross-spectral purity was originally introduced by L. Mandel, *J. Opt. Soc. Am.* **51**, 1961, 1342.
24. L. Mandel and E. Wolf, *J. Opt. Soc. Am.* **66**, 1976, 529.
25. R. Hanbury-Brown and R. Q. Twiss, *Nature* **177**, 1956, 27.
26. R. Barakat, *Opt. Commun.* **23**, 1977, 147.
27. R. Barakat, *Optica Acta* **30**, 1983, 1171.
28. C. Brosseau, *Optik* **104**, 1996, 21.
29. H. Poincaré, *Théorie Mathématique de la Lumière*, Georges Carré, Paris, 1892, Vol. 2.
30. S. Pancharatnam, *Proc. Indian Acad. Sci.* **A44**, 247, 1956.
31. M. Berry, *Proc. Roy. Soc. Lond.* **A392**, 1984, 45. See also A. Shapere and F. Wilczek, eds., *Geometric Phases in Physics*, World Scientific, Singapore, 1989.
32. R. Tycho, *Phys. Rev. Lett.* **58**, 1987, 228.
33. D. Bitter and D. Dublers, *Phys. Rev. Lett.* **59**, 1987, 251.
34. P. K. Aravind, *Opt. Commun.* **94**, 1992, 191.
35. E. De Vito and A. Levrero, *J. Mod. Opt.* **41**, 1994, 233.
36. R. Bhandari and J. Samuel, *Phys. Rev. Lett.* **60**, 1988, 1211. See also J. Samuel and R. Bhandari, *Phys. Rev. Lett.* **60**, 2339, 1988.
37. T. H. Chyba, L. J. Wang, L. Mandel, and R. Simon, *Opt. Lett.* **13**, 1988, 562.
38. L. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms*, Wiley, New York, 1975.
39. V. S. Zapasskii and G. G. Kozlov, *Opt. Spectrosc.* **78**, 1995, 88.
40. The computation of the moments of a complex Gaussian random process has an extensive history. See, for instance, K. S. Miller, *SIAM Rev.* **11**, 1969, 544. See also K. S. Miller, *Complex Stochastic Processes*, Addison-Wesley, Reading, MA, 1974.
41. M. Abramovitz and I. Stegun, *Handbook of Mathematical Functions*, Natl. Bur. Std., U.S. Govt. Printing Office (GPO), Washington, DC, 1968.
42. D. K. C. MacDonald, *Proc. Camb. Phil. Soc.* **45**, 1949, 368; V. I. Bunimovitch, *Zhur. Tekh. Fiz.* **19**, 1949, 1231.
43. B. E. A. Saleh, *J. Opt. Soc. Am.* **63**, 1973, 422 and the comment by R. M. A. Azzam and N. B. Bashara, *J. Opt. Soc. Am.* **64**, 1974, 398.
44. C. Brosseau, E. Rockower, and R. Barakat, *Opt. Commun.* **82**, 1991, 204.
45. The characterization of the temporal statistics of the Stokes parameters have also been discussed by D. Eliyahu, *Phys. Rev.* **E47**, 1993, 2881.
46. A. F. Fercher and P. F. Steeger, *Opt. Acta* **28**, 1981, 443. See also A. F. Fercher and P. F. Steeger, *Opt. Acta* **39**, 1984, 1395; P. F. Steeger, T. Asakura, K. Zocha, and A. F. Fercher, *J. Opt. Soc. Am.* **A1**, 1984, 677 and P. F. Steeger, *Opt. Lett.* **8**, 1983, 528.
47. C. Brosseau, *J. Mod. Opt.* **39**, 1992, 1167.

48. C. Brosseau, *Appl. Opt.* **34**, 1995, 4788. See also V. N. Kurashov, V. V. Marenko, and T. V. Molebnaya, *Opt. Spectrosc. (USSR)* **70**, 1991, 236.
49. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, New York, 1965.
50. V. N. Kurashov, V. V. Marenko, and T. V. Molebnaya, *Opt. Spectrosc. (USSR)* **71**, 1992, 573.
51. A. K. Jaiswal and C. L. Mehta, *Phys. Rev.* **186**, 1969, 1355. See also C. L. Mehta, *Theory of Photoelectric Counting*, in *Progress in Optics*, Vol. VIII, E. Wolf, ed., North Holland Publishing Company, Amsterdam, 1970.
52. G. G. Stokes, *Trans. Camb. Phil. Soc.* **9**, 1852, 399. See also G. G. Stokes, *Cambridge Dublin Math. J.* **4**, 1849, 1 and G. G. Stokes, *Mathematical and Physical Papers*, Cambridge Univ. Press, London, 1901.
53. E. Verdet, *Lecons d'Optique Physique*, Imprimerie Impériale, Paris, 1869. See also E. Verdet, *Ann. Scientif. Ecole Normale Supérieure* **2**, 1865, 291.
54. R. Barakat, *J. Opt. Soc. Am.* **A6**, 1989, 649.
55. V. K. Potekhin, *Opt. Commun.* **46**, 1983, 261.
56. S. Pancharatnam, *Collected Works*, Oxford Univ. Press, Oxford, 1975.
57. W. Wien, *Ann. Physik* **52**, 1894, 132.
58. M. Planck, *Vorlesungen über die Theorie der Warmestrahlung*, 5th ed., Barth, Leipzig, 1923.
59. M. von Laue, *Ann. Phys.* **20**, 1906, 365; *ibid.* **23**, 1907, 1,795.
60. R. Clark Jones, *J. Opt. Soc. Am.* **53**, 1963, 1314.
61. E. L. O'Neill, *Introduction to Statistical Optics*, Addison-Wesley, Reading, MA, 1963.
62. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton Univ. Press, Princeton, 1955. See also John von Neumann, *Collected Works*, A. H. Taub, ed., Macmillan, New York, 1963, Vol. VI.
63. R. Wehrh, *Rev. Mod. Phys.* **50**, 1978, 221.
64. The foundations of MEP (also known as Jaynes' principle) have been discussed in many texts. For a review of some of these works, see, for example, E. T. Jaynes, *Phys. Rev.* **106**, 620 (1957); E. T. Jaynes, in *Papers on Probability, Statistics and Statistical Physics*, R. D. Rosenkrantz, ed., Reidel, Dordrecht, 1983 and *The Maximum Entropy Formalism*, R. D. Levine and M. Tribus, eds., MIT Press, Cambridge, MA, 1979 and J. Skilling and S. I. Gull, "Algorithms and applications," in *Maximum Entropy and Bayesian Methods in Inverse Problems*, C. R. Smith and W. T. Grandy, eds., Kluwer Academic, Dordrecht, 1985.
65. C. Brosseau, *C. R. Acad. Sci. (Paris)* **B310**, 1990, 181. See also C. Brosseau, *Optik* **88**, 1991, 109 and R. Barakat and C. Brosseau, *J. Opt. Soc. Am.* **10**, 1993, 529.
66. C. Brosseau and D. Bicout, *Phys. Rev.* **E50**, 1994, 4997. See also D. Bicout and C. Brosseau, *J. Phys. I. France* **2**, 1992, 2047.
67. L. Landau and L. Lifshitz, *Statistical Physics*, 2nd ed., Pergamon Press, New York, 1977.

## PART 4

---

# INTERACTION OF RADIATION WITH LINEAR MEDIA

When light in an arbitrary state of polarization propagates through a scattering medium or interacts with some optical element, its polarization properties are modified. Broadly speaking, we may categorize the physical mechanisms that can change the state of polarization of light into four areas: (1) propagation through anisotropic media, (2) scattering by a system of particles, (3) oblique reflection and refraction at interfaces, and (4) diffraction by apertures. Here we put most of the emphasis on the first two types of interaction. The effect of non-image-forming optical instruments and scattering media on an incident electromagnetic plane wave is to transform the Jones vector and the Stokes vector so that the medium can be represented by a transformation matrix. All these problems involve sets of linear equations, and matrix methods provide an appropriate and compact way to deal with these problems.

The purpose of this part is to study these transformation matrices, what we term the *Jones matrix* when it acts on the Jones vector and the *Mueller matrix* when it acts on the Stokes vector. The main objective is to help the optical physicist take advantage of these polarization properties by utilizing the information contained in these transformation matrices, which are now routinely measured. The analysis is carried out in a linear theory of the electromagnetic field so that effects such as frequency mixing (e.g., Raman scattering) or multiplication (e.g., second-harmonic generation) are excluded. If the source is a laser, we may need to go beyond this. It is also common to distinguish between passive devices (when light is injected from an external source into the medium) and active devices (where light is generated by the internal gain medium). In the following, we shall be concerned only with passive linear systems. Moreover, we can divide the set of these transformation matrices into two categories: the deterministic and the random. Deterministic optical elements are characterized by transformation matrices whose elements have definite values, while random systems are those for which they are represented by an ensemble of realizations.

Before proceeding, it may be worth to give the detailed plan of this part. In Section 4.1 we discuss the properties of the deterministic Jones and Mueller polarization transfer matrices. Our purpose is to establish the conceptual foundations of these polarization matrix formalisms and show their immediate

implementation. The mathematical relationships between Jones and Mueller approaches to polarization optics are then developed. Among the specific problems we discuss are the conditions of physical realizability of the transformation matrices. The conditions for a Mueller matrix to be derivable from a Jones matrix are also examined. In Section 4.2 we use the method to derive the evolution of polarized light within an optical setup, by the Poincaré sphere method. The utility of this representation is that it provides us with a natural coordinate system in which to analyze the evolution of light polarization in an optical system. Our attention is then turned, in Section 4.3, to an analysis of the polarization effects at dielectric plane interfaces by using the Mueller formalism. The inclusion of reciprocity and symmetry constraints into the analysis of polarization matrices is carried out in Section 4.4. In Section 4.5, we are concerned with the formalism relevant to description of interactions of partially polarized light with a random linear medium. We examine two kinds of randomness. On one hand, we are concerned with an analysis of the effects of temporal fluctuations in a scattering medium on an incident plane quasi-monochromatic wave. On the other hand, we investigate, analytically and by numerical experiments, how multiple scattering of an incident pure state of polarization by a system of independent spherical dielectric particles modifies the polarization state of the probe lightwave. We would like to outline how polarization of an incident wave changes as the wave is scattered inside the material through which it propagates, depending on various physical parameters, including the wavelength of light, the material thickness, and the size of the scatterers. The angular dependence of the light that emerges from the randomly particulate medium has different characteristics than the incident light and carries information about the scattering medium and the underlying propagation mechanisms. Finally, we develop an appropriate formalism to describe the problems of depolarization and enhanced coherent backscattering when light propagates in a medium with randomly positioned isotropic scatterers.

## SECTION 4.1

# Jones and Mueller Polarization Transfer Matrix Methods

At present, popular theories of polarized radiation interaction with optical elements or scattering media may be divided into two groups: the Jones calculus, which assumes a coherent addition of waves; and the Mueller calculus, which assumes an incoherent addition of waves [1, 2]. However, the choice is not a matter of taste. In both approaches, one usually starts a theoretical analysis with severely restrictive assumptions. In the Jones calculus, one starts out with Maxwell's equations, whereas in the Mueller calculus, it starts by postulating a linear relation between the input Stokes vector and the output Stokes vector emerging from the optical medium [3]. The development of the Mueller analysis is heuristic and lacks the mathematical rigor of the Jones calculus. However, the Mueller formalism has an advantage in that it deals with intensities rather than field vectors [4]. The question naturally arises as to whether these two formalisms are equivalent for the description of polarization phenomena. We want to explore that problem here, both because it is an instructive exercise and because it gives some insight about the physical significance of the Jones and Mueller formalisms. As we shall see in detail in the section to follow, the answer to this question depends on the depolarizing character of the optical system.

The problem of propagation through nonimaging optical devices<sup>1</sup> or scattering by optical media may be approached in the frequency or time domains. One can argue that inasmuch as we deal with interactions of quasimonochromatic plane waves in the steady state with polarization-sensitive devices, we must work in the time domain rather than in the spectral domain. From the physical point of view, some experiments are best understood in the frequency domain because their essential physics is of short duration, such as with transient phenomena. It turns out that both analyses are connected by a Fourier transform as reviewed in a seminal work of Parke [7]. More recently, the interrelation between the temporal and spectral representations has been dealt with by several authors [8].

<sup>1</sup>A *nonimaging optical device* consists of polarizers, rotators, and compensators. An *imaging optical system* is composed of lenses and spatial filters.

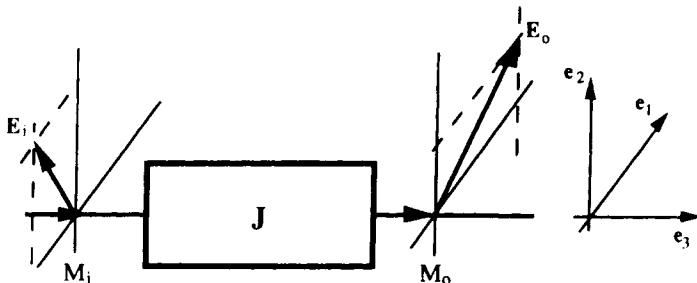
In discussing polarization effects, it is convenient to use normally incident radiation within the plane-wave approximation,<sup>2</sup> with plane perpendicular to the direction of propagation. In the far field, the emerging wave from the optical medium can be treated as a plane wave because usual experimental dimensions are many orders of magnitude larger than, say, wavelength dimension. Most of the energy is transmitted (or scattered) in the specular direction and the paraxial approximation is applicable [9]. We emphasize at the outset that manipulations of these polarization matrices require a precise understanding of the coordinate system involved. We also assume that the optical system is free of internal reflections, with no reflections of light from either surface of the medium.

The remainder of this chapter is subdivided as follows. In the sections that follow we lay out the bare bones of the Jones and Mueller polarization matrix theories and fix the notation and terminology of our description. In Section 4.1.1 the Jones calculus is introduced. Then, Section 4.1.2 provides a basic description of some general features of the Mueller matrix formalism. After some preliminaries, we perform a geometric analysis of optical polarizing elements in terms of parametric equations on the Poincaré sphere. Section 4.1.3 summarizes the various aspects of the problem of the relationship between Jones and Mueller polarization matrices, such as the Mueller–Jones equivalence issue. Finally, in Section 4.1.4 we close with brief remarks on other matrix formalisms. We shall presume here an acquaintance with Part 3 of this book, specifically its Sections 3.1 and 3.3, so that we may avail ourselves of several important definitions introduced here; also, a working knowledge with matrix algebra will afford a helpful perspective on our analysis of Jones and Mueller formalisms.

#### 4.1.1. JONES CALCULUS

Consider first the situation where both input and output light beams are described by pure states ( $P_i = P_o = 1$ ). A flowchart to guide our analysis is sketched in Fig. 4.1.1. When a quasimonochromatic light beam is passed through a linear and deterministic optical device (e.g., scattering medium, optical instrument) such that the input  $\mathbf{E}_i$  and output  $\mathbf{E}_o$  Jones vectors at given points (resp.  $M_i$  and  $M_o$ ) can be treated as plane waves propagating in the same direction characterized by the unit vector  $\mathbf{e}_3$ , then it follows from the linearity of Maxwell's equations that the outgoing field  $\mathbf{E}_o$  has components that are linearly related to those of the incoming field  $\mathbf{E}_i$ . Physically, the optical device alters the incident state of polarization by changing the amplitudes and/or the phases of the components of the Jones vector. The linear transform-

<sup>2</sup>The assumption of plane waves and planar surfaces is adequate for most applications. The case of a plane wave incident on a curved interface has been discussed by Love and Snyder [5]. Regarding further details concerning the polarization effects for spherical incident waves, we refer the reader to Ref. 6.



**Figure 4.1.1.** Input-output representation of the optical system placed in vacuum. Coordinates  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  form a right-hand triplet of unit vectors along the coordinate axes  $(x_1, x_2, x_3)$ . The  $x_1$  and  $x_2$  axes are regarded as the horizontal and vertical directions, respectively. The direction of propagation is along  $x_3$ . The optical system is bounded by two parallel planes that act as input and output planes.

ation of the incident field  $\mathbf{E}_i$  at a point  $M_i$  in the incident plane into the transmitted field  $\mathbf{E}_o$  at a point  $M_o$  in the output plane results from the reradiation of light by oscillating dipoles and currents induced by the incident field within the linear response viewpoint.

With reference to Fig. 4.1.1, we symbolize the optical medium (or instrument) under consideration by an input-output device bounded by two parallel planes.<sup>3</sup>

The linearity of the Maxwell equations allows one to write the relation between the Jones vectors  $\mathbf{E}_i$  and  $\mathbf{E}_o$  by

$$\mathbf{E}_o(t; T) = \mathbf{J} \mathbf{E}_i(t; T) \quad (4.1.1)$$

where we have used Eq. (3.1.10) and  $\mathbf{J}$ , the  $2 \times 2$  Jones matrix of complex-valued elements  $J_{kl}$  ( $k, l = 1, 2$ ):

$$\mathbf{J} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix} \quad (4.1.2)$$

Put another way, the  $\mathbf{J}$  transfer matrix couples the input wave of the system to the output wave of that system, in the paraxial approximation. To avoid a proliferation of symbols, we will drop the argument of the Jones vectors in the following discussion, keeping in mind that all quantities are specified in the time domain.

It is interesting to note that in general, instruments are frequency-dependent, that is, they introduce different effects in different frequency components of the

<sup>3</sup>The  $2 \times 2$  transformation matrices of the Jones calculus are inapplicable to materials having a large reflectance, such as thin films with thickness somewhat smaller than the optical wavelength.

wavefield. Under the quasimonochromatic assumption that the waves form a narrow band about a large central frequency  $v_0$ , all frequency-dependent quantities may be evaluated at  $v_0$ . For a polychromatic radiation field, the transformation of the state of polarization suffered by different Fourier components of the wave will be, in general, different. We also remark that a polychromatic wavefield considered as a superposition of the individual Fourier components is, in general, partially polarized even if each Fourier components is completely polarized.

Before going further, it may be in order to remark on Eq. (4.1.2). The four complex numbers  $J_{kl}$  of the Jones matrix contain eight real parameters, but if we take into account the irrelevance of its overall (nonmeasurable) phase, we are left with just seven parameters. It contains everything that we can expect to find out about the optical medium. More specifically, these parameters are related to eight independent elementary properties of the medium: isotropic absorption of the amplitude and isotropic retardation of the phase of the wave, linear birefringence and linear dichroism along the  $x_1-x_2$  axes, linear birefringence and linear dichroism along the bisectors of the coordinate axes, and circular birefringence and circular dichroism. It should be noted that Clark Jones was evidently the first to derive the matrix for the general case of a medium simultaneously requiring all eight types of canonical behavior [3]. We shall return to this aspect in a more advanced form below in Section 5.1.

Implicit in the foregoing definition of the Jones matrix is a choice of an orthonormal basis. The four components of  $\mathbf{J}$  are all defined with respect to this basis. The orthogonal basis states of polarization that are used to define the Jones matrix may be considered arbitrary. In general, we assume the orthogonal 1 and 2 linear polarizations to be the basis states. One can ask how these components would change if we were to switch to a new basis for the Jones vector, say,  $(\mathbf{e}_u, \mathbf{e}_v)$ . If we set  $\mathbf{E}^{(uv)} = \mathbf{U}\mathbf{J}\mathbf{U}^{-1}$ , where  $\mathbf{U}$  is nonsingular (and therefore having an inverse  $\mathbf{U}^{-1}$ ), and substitute into Eq. (4.1.2), we find that Eq. (4.1.2) will be maintained and only its representation changed if  $\mathbf{J}$  is transformed into

$$\mathbf{J}^{(uv)} = \mathbf{U}\mathbf{J}\mathbf{U}^{-1} \quad (4.1.3)$$

We can now obtain a number of interesting properties from the similarity transformation Eq. (4.1.3). In particular, as the trace and determinant are preserved through a similarity transformation,  $\mathbf{J}^{(uv)}$  and  $\mathbf{J}$ , have the same eigenvalues. For example, the circular Jones matrix is expressed in the helicity basis  $(\mathbf{e}_l, \mathbf{e}_r)$  as

$$\mathbf{J}^{(lr)} = \mathbf{U}_c \mathbf{J} \mathbf{U}_c^{-1} = \begin{bmatrix} J_{11} + J_{22} - i(J_{12} - J_{21}) & J_{11} - J_{22} + i(J_{12} + J_{21}) \\ J_{11} - J_{22} - i(J_{12} + J_{21}) & J_{11} + J_{22} + i(J_{12} - J_{21}) \end{bmatrix} \quad (4.1.4)$$

where  $\mathbf{U}_c$  is as given by Eq. (3.1.20a).

At this point, three additional points may be made. First, the fact that the relationship between the input and output Jones vectors is expressible in the form given by Eq. (4.1.1) does not imply that the most general relation between input and output Jones vectors is consistent with this linear transformation. Reference is made to the operation of complex conjugation, which produces the transformation  $\mathbf{E} \rightarrow \mathbf{E}^*$  for a physical situation where the input and output Jones vectors are not related according to Eq. (4.1.1) [10]. This antiunitary transformation implies, for instance, that a right-hand circularly polarized wave is turned into a left-hand circularly polarized wave. However, a linearly polarized wave remains linearly polarized. The physical device that realizes this operation is the phase conjugation mirror [11]. Second, it may be worth remarking that  $\mathbf{J}$  would be a  $3 \times 3$  matrix if we assume a nonplane wavefield described by a Jones vector that is a  $3 \times 1$  column vector. The third and most crucial point in the discussion is that the Jones formalism was basically tailored to map pure states into pure states, or put in another way, to describe optical systems that do not depolarize pure states of polarization passing through it. However, combined with the coherency matrix formalism, it can be extended to handle transformation of mixed states. We will demonstrate this shortly.

#### 4.1.1.1. Eigenvalue–Eigenvector Decomposition

In general, optical devices divide an incident optical field into two parts, the eigenvalues of the optical device, and transmits these parts with different transmittances and different phases. The Jones matrix formalism takes a simple form through its principal eigenvector–eigenvalue representation. The principal states (eigenpolarizations) are the polarization states that are simply scaled out by the optical element, and they are associated with the eigenvectors of the corresponding Jones matrix. In general, a Jones matrix has two different eigenpolarizations that obey the eigenequations

$$\mathbf{JE}_a = \lambda_a \mathbf{E}_a \quad (4.1.5a)$$

$$\mathbf{JE}_b = \lambda_b \mathbf{E}_b \quad (4.1.5b)$$

where  $\lambda_j = |\lambda_j| \exp(i \arg(\lambda_j))$   $j = a, b$  are the two associated complex eigenvalues. To understand the operation of the optical element described by Jones matrix  $\mathbf{J}$ , the physical picture to imagine is that it modifies the polarization state of the incident light by altering the amplitude ( $|\lambda_a|$ ) and phase ( $\arg(\lambda_a)$ ) of its eigenpolarization independently.

An optical element that is classified as homogeneous has two orthogonal eigenpolarizations (e.g., ideal plane mirror, compensator):  $\mathbf{E}_a$  and  $\mathbf{E}_b$ , which verify Eq. (3.1.47). Inhomogeneous optical elements have nonorthogonal eigenpolarizations (e.g., circular polarizer). The amplitude of the eigenvalue is the amplitude attenuation of the light as it passes through the device in the corresponding eigenvector and the phase of the eigenvalue is the phase delay seen

in the corresponding eigenvector. The parameter  $d = (|\lambda_a|^2 - |\lambda_b|^2)/(|\lambda_a|^2 + |\lambda_b|^2)$  is called the *diattenuation* of the optical element. The difference of phase between the eigenvalues,  $r = |\arg(\lambda_a) - \arg(\lambda_b)|$  with  $0 \leq r \leq \pi$ , is called the *retardation* of the optical element [12].

Denoting by  $\mathbf{E}_a, \mathbf{E}_b$  the two eigenvectors, the expression for the Jones matrix is given by

$$\mathbf{J} = \begin{bmatrix} E_{a1} & E_{b1} \\ E_{a2} & E_{b1} \end{bmatrix} \begin{bmatrix} |\lambda_a| \exp\left(\frac{ir}{2}\right) & 0 \\ 0 & |\lambda_b| \exp\left(-\frac{ir}{2}\right) \end{bmatrix} \begin{bmatrix} E_{a1} & E_{b1} \\ E_{a2} & E_{b1} \end{bmatrix}^{-1} \quad (4.1.6)$$

If the two eigenvectors are orthogonal, then the matrix

$$\begin{bmatrix} E_{a1} & E_{b1} \\ E_{a2} & E_{b1} \end{bmatrix}$$

is unitary. Note that if the two eigenvalues are equal, the Jones matrix is diagonal, proportional to  $\sigma_0$ .

This eigenvalue–eigenvector decomposition has a number of applications. For example, the nature and eigenstates of a laser can be determined from the “resonance” condition, namely,  $\mathbf{J}_l \mathbf{E} = \lambda \mathbf{E}$ , where  $\mathbf{J}_l$  is the Jones matrix for one round trip inside the cavity [13].

#### 4.1.1.2. Jones Formalism and the $2 \times 2$ Coherency Matrix Formulation

For present purpose, we first derive the law of transformation of the coherency matrix. Taking the direct product of each side of Eq. (4.1.1) times its Hermitian adjoint followed by ensemble averaging yields

$$\langle \mathbf{E}_o \otimes \mathbf{E}_o^+ \rangle = \Phi_o = \mathbf{J} \langle \mathbf{E}_i \otimes \mathbf{E}_i^+ \rangle \mathbf{J}^+ = \mathbf{J} \Phi_i \mathbf{J}^+ \quad (4.1.7)$$

where  $\Phi_i$  and  $\Phi_o$  denote respectively the input and output coherency matrices. The congruency transformation [Eq. (4.1.7)] represents the important result of this section and is the key equation for the ensuing analysis. It provides the basic law of interaction of a partially polarized wavefield with a deterministic linear optical device. This equation parallels the well-known interference formula in the passive linear filter theory [14].

Let us examine Eq. (4.1.7) more closely. First, note that if we consider the special case of incident unpolarized light, Eq. (4.1.7) may be rewritten as

$$\Phi_{ou} = \frac{\langle S_0 \rangle_u}{2} \mathbf{J} \mathbf{J}^+ \quad (4.1.8)$$

Second, taking the determinant of both sides of Eq. (4.1.7) yields

$$\det(\Phi_o) = \det(\Phi_i) |\det(\mathbf{J})|^2 \quad (4.1.9)$$

Thanks to Eq. (3.1.147), Eq. (4.1.19) can be rewritten as

$$\langle\langle S\rangle_{o0}^2 - \langle S\rangle_{o1}^2 - \langle S\rangle_{o2}^2 - \langle S\rangle_{o3}^2\rangle = |\det(\mathbf{J})|^2 (\langle S\rangle_{i0}^2 - \langle S\rangle_{i1}^2 - \langle S\rangle_{i2}^2 - \langle S\rangle_{i3}^2) \quad (4.1.10)$$

which implies, making use of Eq. (3.1.154), that

$$(1 - P_o^2) = \frac{|\det(\mathbf{J})|^2}{g^2} (1 - P_i^2) \quad (4.1.11)$$

where we have anticipated the definition of the gain  $g$ , which we will be given shortly.

The usefulness of this expression lies in the fact that under the congruent transformation (4.1.7) an initially polarized beam remains fully polarized, although the form of the polarization state can be altered. The most crucial point in this discussion is that the Jones formalism can be applied to only optical systems that cannot depolarize an incident pure state. Generalization of Eq. (4.1.11) may be derived easily for the  $N$ -fold problem.<sup>4</sup> However, we emphasize strongly that the output degree of polarization of an incident partially polarized polarization state can increase or decrease on transmission through the Jones medium. In the special case of a completely polarized light, we readily find

$$\det(\Phi_{op}) = 0 \quad (4.1.12)$$

This formula ensures that the two-level decomposition of a partially polarized light into totally polarized and unpolarized components is unique. We shall shortly have occasion to make use of these properties.

#### 4.1.1.3. Gain

A result that will prove of great importance in what follows is the expression of the gain of the medium. The gain  $g$  (also termed the *transmittance*) is defined as the output intensity divided by the input intensity

$$g \equiv \frac{\langle S_0 \rangle_o}{\langle S_0 \rangle_i} \quad (4.1.13)$$

<sup>4</sup>The generalization is served through the congruency transformation  $\mathbf{D}_{No} = 1/(g)\mathbf{J}_N\mathbf{D}_{Ni}\mathbf{J}_N^+$ , where  $\mathbf{D}_{Ni}$  (resp.  $\mathbf{D}_{No}$ ) is the input (resp. output) density matrix,  $\mathbf{J}_N$  defines the generalized  $N \times N$  Jones complex-valued matrix and  $g$  is the gain. For  $N = 2$ , we arrive at Eq. (4.1.11) and for  $N = 3$ , the final result is  $1 - (P_{30}^{(3)})^2 = (1 - (P_{31}^{(2)})^2)[|\det(\mathbf{J}_3)|^2/g^3]$ . Notice that for  $N = 3$ , an incident pure state of polarization is not generally in the form of a pure state after interaction with the optical medium since  $\text{tr}(\mathbf{D}^2)$  is not generally conserved, that is,  $\text{tr}(\mathbf{D}_o^2) = \text{tr}[(\mathbf{J}_3^+ \mathbf{J}_3 \mathbf{D}_i)^2] \neq \text{tr}(\mathbf{D}_i^2)$ .

For passive optical systems, it follows that  $0 \leq g \leq 1$ . We start from the fact that everything knowable about a stochastic process can, in principle, be calculated from the knowledge of the density operator. On substituting Eqs. (3.1.128) and (4.1.7) into Eq. (4.1.13), we get

$$g = \text{tr}(\mathbf{D}_2 \mathbf{J}^+ \mathbf{J}) = \langle \mathbf{J}^+ \mathbf{J} \rangle \quad (4.1.14)$$

This equation shows that the gain is related to the expectation value of  $\mathbf{J}^+ \mathbf{J}$ . Note that the expression of  $g$  is basis-independent since  $\mathbf{J}^+ \mathbf{J}$  is independent of what orthonormal basis is used. From a mathematical standpoint we can rephrase this result by stating that the gain is invariant under the change-of-basis unitary transformation.

To produce a more useful relation, we can use Eq. (3.1.144) and substitute into Eq. (4.1.14). By doing so, we find that

$$g = (1 - P)g_u + Pg_p = \left( \frac{1 - P}{2} \right) \text{tr}(\mathbf{J}^+ \mathbf{J}) + Pg_p \quad (4.1.15)$$

This formula reflects the convexity property of  $\mathbf{D}_2$ . Now making use of the Pauli decomposition, it follows at once that  $g$  takes the form

$$g = b_0 + \sum_{k=1}^3 b_k \langle \sigma_k \rangle_i \quad (4.1.16)$$

where  $b_k = \frac{1}{2} \text{tr}(\mathbf{J}^+ \mathbf{J} \sigma_k)$  and  $\langle \sigma_k \rangle_i$  are the input normalized Stokes parameters. From Eq. (4.1.16), one finds how  $g$  depends on the state of polarization of the light incident on the optical system. The expansion coefficients  $b_k$  are given explicitly by

$$b_0 = \frac{1}{2}(|J_{11}|^2 + |J_{12}|^2 + |J_{21}|^2 + |J_{22}|^2) \quad (4.1.17a)$$

$$b_1 = \frac{1}{2}(|J_{11}|^2 - |J_{12}|^2 + |J_{21}|^2 - |J_{22}|^2) \quad (4.1.17b)$$

$$b_2 = \text{Re}(J_{11}^* J_{12} + J_{21}^* J_{22}) \quad (4.1.17c)$$

$$b_3 = \text{Im}(J_{11}^* J_{12} + J_{21}^* J_{22}) \quad (4.1.17d)$$

The gain varies between the limits  $g_{\max}$  and  $g_{\min}$ , which denote respectively the maximum and minimum gains. These limits are most easily determined by observing that  $\sum_k b_k \langle \sigma_k \rangle$  is the inner product of the vector

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

and of the Poincaré vector

$$\begin{bmatrix} \langle \sigma_1 \rangle \\ \langle \sigma_2 \rangle \\ \langle \sigma_3 \rangle \end{bmatrix}$$

we see that  $g$  assumes its minimum (resp. maximum) possible value when these vectors are antiparallel (resp. parallel). It is interesting to note that  $g_{\max}$  and  $g_{\min}$  reduce to the forms

$$g_{\max} = b_0 + P_i \left( \sum_{k=1}^3 b_k^2 \right)^{1/2} \quad (4.1.18a)$$

$$g_{\min} = b_0 - P_i \left( \sum_{k=1}^3 b_k^2 \right)^{1/2} \quad (4.1.18b)$$

We next note that  $g_{\max}$  and  $g_{\min}$  are squares of the singular values of the Jones matrix  $\mathbf{J}$ , that is, they are the eigenvalues of  $\mathbf{JJ}^+$ . Physically it implies the existence of two orthogonal incident states  $\mathbf{E}_{\max}$  and  $\mathbf{E}_{\min}$  corresponding to the maximum and minimum gains. It follows from Appendix I that  $\mathbf{E}_{\max}$  and  $\mathbf{E}_{\min}$  are eigenvectors of  $\mathbf{JJ}^+$ . Moreover, these states satisfy the orthogonality condition

$$\mathbf{E}_{\min}^* \cdot \mathbf{E}_{\max} = \mathbf{E}_{\max} \cdot \mathbf{E}_{\min}^* = 0 \quad (4.1.19)$$

Now an important property of Hermitian matrices is that they can be diagonalized by a unitary transformation. Hence there exists a basis in which  $\mathbf{JJ}^+$  is diagonal. Using this matrix algebra result, we may write  $\mathbf{J}^+ \mathbf{J}$  in the form

$$\mathbf{J}^+ \mathbf{J} = \mathbf{UCU}^+ \quad (4.1.20)$$

where

$$\mathbf{C} = \begin{bmatrix} c_1 & 0 \\ 0 & c_2 \end{bmatrix}$$

and  $\mathbf{U}$  is unitary. Substituting this expansion into Eq. (4.1.14), we get

$$g = \text{tr}(\mathbf{D}_2 \mathbf{UCU}^{-1}) = \text{tr}(\mathbf{CU}^{-1} \mathbf{D}_2 \mathbf{U}) = \text{tr}(\mathbf{CD}_2^{(ab)}) \quad (4.1.21)$$

where  $\mathbf{D}_2^{(ab)}$  is the density matrix written in the orthonormal basis  $(ab)$ . It follows from Eq. (4.1.21) that the formula for the gain reads as

$$g = \frac{(c_1 \langle |E_a|^2 \rangle + c_2 \langle |E_b|^2 \rangle)}{\langle |E_a|^2 \rangle + \langle |E_b|^2 \rangle} \quad (4.1.22)$$

By virtue of the definition of the normalized Stokes parameters  $\langle \sigma_j \rangle^{ab} \equiv \langle S_j \rangle^{ab} / \langle S_0 \rangle^{ab}$ , Eq. (4.1.22) can be rewritten

$$g = \frac{c_1}{2}(1 + \langle \sigma_1 \rangle^{ab}) + \frac{c_2}{2}(1 - \langle \sigma_1 \rangle^{ab}) = \left( \frac{c_1 + c_2}{2} \right) \left( 1 + \left( \frac{c_1 - c_2}{c_1 + c_2} \right) \langle \sigma_1 \rangle^{ab} \right) \quad (4.1.23)$$

this equation shows the dependence of  $g$  on the only relevant parameter  $\langle \sigma_1 \rangle^{ab}$ . A Jones vector  $\mathbf{E}_i$  incident on an optical device that is one of the two eigenpolarizations,  $\mathbf{E}_a$  or  $\mathbf{E}_b$ , is transmitted in the same state. If the incident Jones vector  $\mathbf{E}_i$  is an arbitrary state of polarization, the transmitted state will be different from  $\mathbf{E}_i$ .

Now we observe from Eqs. (4.1.18) that the gain for an input unpolarized light may be taken to be the arithmetic mean of  $g_{\max}$  and  $g_{\min}$

$$g_u = \frac{g_{\max} + g_{\min}}{2} = b_0 = \frac{1}{2} \operatorname{tr}(\mathbf{J}^+ \mathbf{J}) \quad (4.1.24)$$

A final relation follows as a consequence of Eqs. (4.1.18a, b)

$$\frac{g_{\max} - g_{\min}}{g_{\max} + g_{\min}} = P_i \frac{(\sum_{k=1}^3 b_k^2)^{1/2}}{b_0} \quad (4.1.25)$$

This formula is important in the context of interferometry since the left-hand side of Eq. (4.1.25) defines the visibility of fringes at a given point in the fringe pattern. It is also worth noting that this expression is invariant with respect to rotations of the axes.

Practical applications of the preceding equations are numerous. The first term of the right-hand side of Eq. (4.1.16) depends only on  $\mathbf{J}$ , while the second term is linear in the polarization characteristics. We will return to this point later, in Part 5, since it is the basis of an important method (differential polarized spectroscopy) of measurement of the Jones matrix elements [15]. For practical purpose, it is also appropriate to emphasize the importance of the absorbance  $A \equiv -\ln(g)$ . This is due to the multiplicative character of the gain through different media in series, which becomes extensive in the  $\ln(g)$  formulation.

In closing this section we would like to point out that if the Jones matrix  $\mathbf{J}$  is unitary (i.e.,  $\mathbf{J}^+ \mathbf{J} = \sigma_0$ ) and nonsingular, then  $\det(\mathbf{D}_2)$ ,  $\operatorname{tr}(\mathbf{D}_2)$ , and  $\operatorname{tr}(\mathbf{D}_2 \ln(\mathbf{D}_2))$  are similarity invariants. Such an optical medium is invariant with respect to intensity, degree of polarization, and entropy, although it can alter the form of polarization. Examples of unitary Jones matrices include compensators and rotators. The special case of singular Jones matrices [i.e.,  $\det(\mathbf{J}) = 0$ ] corresponds to the perfect polarizers.

#### 4.1.1.4. Jones Matrices for Common Optical Devices

We next develop the matrix representations of certain common optical components. In particular, we show that interesting group theory tricks can be used to make the analysis possible in an efficient way. The analysis about to follow is based on the invariants under similarity transformations. Group theory provides the quickest and most elegant way of identifying these invariant features. These invariants are the trace, the determinant, and the rank. The rank of a square matrix is the number of linearly independent columns (or rows). Now  $2 \times 2$  matrices can have rank 0, 1, or 2. Excluding the null matrix of rank 0, we have, on one hand, Jones matrices of rank 1 matrices that are such that  $\det(\mathbf{J}) = 0$  and  $\text{tr}(\mathbf{J}) \neq 0$ . On the other hand, Jones matrices of rank 2 are such that  $\det(\mathbf{J}) \neq 0$  and  $\text{tr}(\mathbf{J}) \neq 0$ , that is, they are not singular and therefore possess an inverse. In view of these remarks, we define three families of Jones matrices:

1. The singular transformations,  $\det(\mathbf{J}) = 0$ . Singular Jones matrices constitute a five-parameter family. Such matrices are called *polarizers* because the output degree of polarization  $P_o = 1$ , regardless of the input state of polarization.
2. The nonsingular transformations, which constitute a seven-parameter family. It is important to appreciate that the unit matrix  $\sigma_0$  falls into this category and that each such matrix possesses an inverse — the nonsingular Jones matrices form a group with respect to the multiplication.
3. The unitary transformations, which introduce only a phase shift and then have the following parametrization as the imaginary part exponential of a Hermitian matrix (i.e.,  $\mathbf{L}^+ = \mathbf{L}$ ):

$$\mathbf{J} = \exp(i\mathbf{L}) \quad (4.1.26)$$

Unitary Jones media have a constant gain equals to one independent of the incident state of polarization. The set of Jones matrices that let invariant the gain constitutes the group  $U_2$ . As  $\mathbf{J}$  is a unitary matrix, we have  $\det(\mathbf{J}) = \exp(i\phi)$ , and we can write  $\mathbf{J} = \exp(i\phi/2)\mathbf{V}$ , which is equivalent to Eq. (4.1.26). In this expression  $\mathbf{V}$  is a unitary and unimodular matrix, specifically,  $\det(\mathbf{V}) = 1$ . The Jones matrices of this form constitute a four-parameter family. The group of unitary and unimodular  $2 \times 2$  matrices  $SU(2)$  is isomorph with the group of rotations of the three-dimensional real space,  $O(3)$ . The mapping is two to one involving double angles in the real 3-space. All transformations of this group can be expanded in terms of three independent generators  $\mathbf{J}_k$  ( $k = 1, 2, 3$ ) with

$$\mathbf{J}_k = \cos(\theta_k)\sigma_0 + i \sin(\theta_k)\sigma_k \quad (4.1.27)$$

where  $\theta_k$  is a phase angle.

**TABLE 4.1.1. Jones Matrices for Various Optical Devices and Their Relevant Characteristics**

Jones Matrix of Optical Device	Gain	$g_{\max}$ $g_{\min}$	Eigenvalues $\lambda_a, \lambda_b$	Normalized Eigenpolarization $E_a, E_b$
$P(\alpha)$	$\frac{1}{2}(1 + \cos(2\alpha)\langle\sigma_1\rangle + \sin(2\alpha)\langle\sigma_2\rangle)$	1 0	0 1	$[\sin(\theta), -\cos(\theta)]^T$ $[-\cos(\theta), \sin(\theta)]^T$
$R(\theta)$	1		$\exp(i\theta)$ $\exp(-i\theta)$	$[1, i]^T$ $[1, -i]^T$
$C(\delta)$	1		$\exp\left(i\frac{\delta}{2}\right)$ $\exp\left(-i\frac{\delta}{2}\right)$	$[1, 0]^T$ $[0, 1]^T$

With these remarks, we now outline a derivation of the explicit forms of Jones matrices for various canonical optical devices. Observe that all quantities are defined in the linear polarization basis, unless otherwise specified. As we have remarked above, any Jones matrix is determined up to the complex factor. As an aid to the reader, we summarize the relevant characteristics of these optical components in Table 4.1.1.

A typical example of Jones optical medium of the family 2 in the preceding list is the polarizer. Depending on the form of the eigenpolarizations, a polarizer is classified as linear, circular, or elliptical. A linear polarizer has linearly polarized eigenpolarizations, whereas a circular (resp. elliptical) polarizer has circularly (resp. elliptically) polarized eigenpolarizations. A polarizer changes only the amplitudes of the components of the incident Jones vector. An ideal linear polarizer is an optical system that passes only the component of the electric field making an angle  $\alpha$  with respect to the 1 direction (transmission axis<sup>5</sup>) for arbitrary incident polarization states. Many devices have been proposed and used as linear polarizers. The most common types of linear polarizer are based either on birefringence (i.e., different light retardation for orthogonal polarization directions) or dichroism (i.e., selective absorption of one polarization component). Typical examples include the dichroic polarizer, such as Polaroid filters, the dye polarizer, and the Glan–Thomson prism. They will be discussed in more detail in Section 5.2. The Jones matrix for this

<sup>5</sup>The transmission axis of a linear polarizer is defined with respect to a linearly polarized light beam normally incident on the face of the polarizer. The transmission axis defines the direction that the Jones vector of the wave must have for the actual gain to be maximized; thus the light emerging from the polarizing filter is linearly polarized along the transmission axis.

device is

$$\mathbf{P}(\alpha) = \begin{bmatrix} \cos^2(\alpha) & \cos(\alpha) \sin(\alpha) \\ \cos(\alpha) \sin(\alpha) & \sin^2(\alpha) \end{bmatrix} = \frac{1}{2}(\sigma_0 + \sin(2\alpha)\sigma_1 + \cos(2\alpha)\sigma_2) \quad (4.1.28)$$

There are several interesting features of Eq. (4.1.28). For instance, the Jones matrix of an ideal linear horizontal polarizer (transmission axis in the 1 direction) is simply

$$\mathbf{P}(0) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad (4.1.29)$$

Thus the output Jones vector is always linearly horizontally polarized regardless of the incident Jones vector:

$$\begin{bmatrix} E_1 \\ E_2 \end{bmatrix} \rightarrow \begin{bmatrix} E_1 \\ 0 \end{bmatrix} = E_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

The corresponding Jones matrix of an ideal linear vertical polarizer is obtainable in a manner identical to that leading to (4.1.29), and we get

$$\mathbf{P}\left(\frac{\pi}{2}\right) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad (4.1.30)$$

Several properties are of interest to us. First, the action of rotating a linear polarizer through an angle  $\alpha$  with respect to the 1 direction yields

$$\mathbf{P}(\alpha) = \mathbf{R}(-\alpha)\mathbf{P}(0)\mathbf{R}(\alpha) \quad (4.1.31)$$

To prove this statement, we consider the rotation matrix

$$\mathbf{R}(\alpha) = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

This permits us to write

$$\begin{aligned} \mathbf{P}(\alpha) &= \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \\ &= \begin{bmatrix} \cos^2(\alpha) & \cos(\alpha) \sin(\alpha) \\ \cos(\alpha) \sin(\alpha) & \sin^2(\alpha) \end{bmatrix} \end{aligned} \quad (4.1.32)$$

which is what we wanted to prove. Mathematically,  $\mathbf{P}(\alpha)$  is called a *projection matrix* in linear algebra. As a consequence of the idempotency of projectors,

further application of the same polarizer has no further effect:

$$(\mathbf{P}(\alpha))^N = \mathbf{P}(\alpha) \quad (4.1.33)$$

This matrix has an additional important property, namely, the orthogonality relation

$$\mathbf{P}(\alpha)\mathbf{P}\left(\alpha + \frac{\pi}{2}\right) \equiv 0 \quad (4.1.34)$$

In addition to its mathematical significance, this equation also has an experimental consequence that can be easily checked. The gain of the system composed of two ideal crossed polarizers, having orthogonal transmission axes, is zero.

On substitution of Eq. (4.1.28) into Eq. (4.1.16), the gain becomes

$$g = \frac{1}{2} + \frac{\cos(2\alpha)}{2} \langle \sigma_1 \rangle + \frac{\sin(2\alpha)}{2} \langle \sigma_2 \rangle \quad (4.1.35)$$

In general, the gain depends on the incident state of polarization. However, the result in Eq. (4.1.35) shows that the output intensity leaving the polarizer is independent of  $\alpha$  if the incident state of polarization is either unpolarized or (totally or partially) circularly polarized. Note that Eq. (4.1.35) also reduces to

$$g = \frac{1}{2}(1 + P_i \cos(2\chi_i) \cos(2(\alpha - \psi_i))) \quad (4.1.36)$$

if we substitute Eq. (3.1.57) into Eq. (4.1.34). The gain, for an incident partially polarized light, varies according to Eq. (4.1.35), that is, it decreases to a minimum  $(1 - P_i)/2$  (for  $\alpha = 0$ ) and then increases sinusoidally to a maximum  $(1 + P_i)/2$  (for  $\alpha = \pi/2$ ) if the polarizer is rotated (sometimes termed the *gain sinusoid*). The gain of an ideal polarizer for an input unpolarized light is a constant equal to

$$g_u = \frac{1}{2} \quad (4.1.37)$$

and is reminiscent to the two-level decomposition, by virtue of Eq. (3.1.144). In practice the existing polarizers are nonideal, that is, common polarizers transmit less than 50% of unpolarized incident light since some light is absorbed in its bulk and some is reflected at its surfaces; moreover, when two such polarizers are crossed, the gain is nonzero [16].

In view of what was said above, the eigenpolarizations describe its principal axes. The orientation corresponding to  $g_{\min}$  and  $g_{\max}$  are orthogonal to one another and are termed the *principal component polarization directions*.

We can extend the analysis to include the case of ideal circular polarizers: left circular polarizer

$$\mathbf{P}_{cl} = \frac{1}{2} \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \quad (4.1.38a)$$

and right circular polarizer

$$\mathbf{P}_{cr} = \frac{1}{2} \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix} \quad (4.1.38b)$$

Regardless of the state of polarization of the incident light, the output state of polarization is always circularly polarized. We leave to the reader the task of proving that the connection between these two sets of polarizers is given by

$$\mathbf{U}_c^{-1} \mathbf{P}(0) \mathbf{U}_c = \mathbf{P}_{cr} \quad (4.1.39a)$$

$$\mathbf{U}_c^{-1} \mathbf{P}\left(\frac{\pi}{2}\right) \mathbf{U}_c = \mathbf{P}_{cl} \quad (4.1.39b)$$

where the constant matrix  $\mathbf{U}_c$  is given by Eq. (3.1.20a).

Typical examples of Jones media of family 3 in the preceding list are the rotator and the compensator. Again restricting our consideration to ideal systems, we define a *rotator* as an optical device that has for effect to produce a rotation of the orthogonal components of the electric field as it propagates through the element. At the output of the rotator the orthogonal components of the electric field, we may write

$$\begin{aligned} E'_1 &= E_1 \cos(\theta) + E_2 \sin(\theta) \\ E'_2 &= -E_1 \sin(\theta) + E_2 \cos(\theta) \end{aligned} \quad (4.1.40)$$

where  $\theta$  denotes the angle through which the field is rotated. The Jones matrix for a rotator is obviously represented by a unitary rotation matrix defined in the usual way for the linear-polarization-state basis set

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} = \sigma_0 \cos(\theta) + i\sigma_3 \sin(\theta) \quad (4.1.41)$$

We note some familiar results:

$$\mathbf{R}^{-1}(\theta) = \mathbf{R}(-\theta) \quad (4.1.42a)$$

$$\mathbf{R}(\theta_1 + \theta_2) = \mathbf{R}(\theta_1)\mathbf{R}(\theta_2) = \mathbf{R}(\theta_2)\mathbf{R}(\theta_1) \quad (4.1.42b)$$

We have seen that, in the definition of a Jones matrix, a choice of an

orthonormal basis must be made. If the helicity basis is adopted, then the circular Jones matrix of an optical rotator takes the diagonal form

$$\mathbf{R}^{(lr)}(\theta) = \begin{bmatrix} \exp(-i\theta) & 0 \\ 0 & \exp(i\theta) \end{bmatrix} \quad (4.1.43)$$

Note that the matrix  $\sigma_3$  is a rotator that rotates the Jones vector through an angle  $\pi/2$ :

$$\mathbf{R}\left(\frac{\pi}{2}\right) = i\sigma_3 \quad (4.1.44)$$

We observe that the relation  $\sigma_3^2 = \sigma_0$  also takes a simple meaning—a double application of the operator  $\sigma_3$  that induces a rotation through  $\pi$  is equivalent to the operation of the unit operator,  $\mathbf{R}^2(\pi/2) = -\sigma_3^2 = -\sigma_0$ . Note that the rotation induced by an optically active medium may be described by a Jones matrix  $\mathbf{R}(\theta)$  with  $\theta$  depending on frequency, namely, Cauchy's formula  $\theta(v) = A + Bv^2$ , where the constants  $A$  and  $B$  are tabulated [17]. The Faraday rotator consists of a suitable medium, such as quartz, placed within a magnetic field coil. The external magnetic field  $\mathbf{H}$  applied to this medium rotates the state of polarization of light passing through the medium. Denoting by  $L$  the path traversed by light in the optical medium, the angle of rotation is defined by the phenomenologic expression  $\theta = V \int_L \mathbf{H} \cdot d\mathbf{l}$  where  $V$ , termed the *Verdet constant*, is a constant that is a property of the material. The Verdet constant varies with both frequency and temperature [11].

An ideal compensator with small retardation  $\delta \ll v_o/\Delta v$  is an optical device that introduces a phase change  $\delta_1$  in the 1 component and  $\delta_2$  in the 2 component for each spectral component of the electric field vector, resulting in a differential phase shift  $\delta = \delta_1 - \delta_2$  between the two orthogonally linearly polarized components of the optical field. Note that the words compensator, homogeneous linear retarder and phase shifter are used synonymously. Before we can proceed any further, we must consider the "fast" and "slow" eigenwaves of the compensator. The fast and slow states of the compensator are given by the eigenpolarizations and propagate with their own phase velocities. The directions of polarization for these eigenstates are mutually orthogonal and define the fast and slow axes of the compensator for that direction of propagation. Because of the difference in phase velocity, one component is retarded with respect to the other. The fast axis is defined to be along the eigenpolarization that emerges first from the compensator, that is, the eigenpolarization with the leading phase.

To develop the theory we consider a normally incident wave characterized by Jones vector

$$\mathbf{E}_i = \begin{bmatrix} E_{i1} \\ E_{i2} \end{bmatrix}$$

The Jones vector of the wave after passing through the compensator takes the form

$$\begin{aligned} E_{o1} &= E_{i1} \exp\left(i \frac{\delta}{2}\right) \\ E_{o2} &= E_{i2} \exp\left(-i \frac{\delta}{2}\right) \end{aligned} \quad (4.1.45)$$

If we assume that  $\delta$  is frequency-independent, then the Jones matrix of a compensator becomes

$$\mathbf{C}(\delta) = \begin{bmatrix} \exp\left(i \frac{\delta}{2}\right) & 0 \\ 0 & \exp\left(-i \frac{\delta}{2}\right) \end{bmatrix} = \sigma_0 \cos\left(\frac{\delta}{2}\right) + i \sigma_1 \sin\left(\frac{\delta}{2}\right) \quad (4.1.46)$$

Conventional phase retardation plates are either natural or induced linear birefringence in a slab of transparent crystals. Refer to Section 5.2 for details of practical importance. For a normal incident radiation, the phase retardation introduced by the waveplate is proportional to the birefringence of the crystal

$$\delta \equiv \frac{2\pi v_0 L \Delta n}{c} \quad (4.1.47)$$

Here  $L$  and  $\Delta n = n_e - n_o$  denote the path length through the crystal and the linear birefringence of the compensator where  $n_e, n_o$  are the refractive indices of the fast and slow waves, respectively.<sup>6</sup>

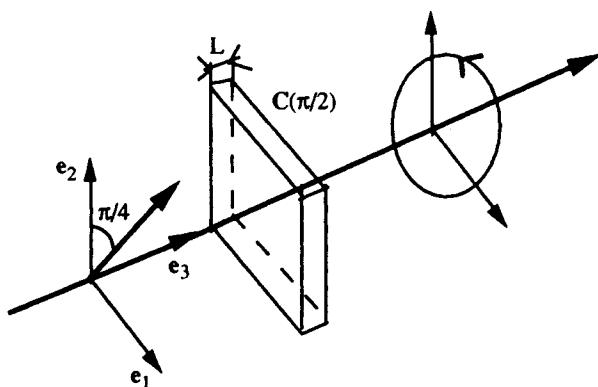
Two cases are of special consideration. First the Jones matrix of a quarter-waveplate is given by

$$\mathbf{C}\left(\frac{\pi}{2}\right) = \begin{bmatrix} 1 & 0 \\ 0 & -i \end{bmatrix} \quad (4.1.48)$$

where we have ignored a  $\pi/4$  phase term. The 2 component is retarded in phase by  $\pi/2$  relative to the 1 component, yielding an output Jones vector that is right

<sup>6</sup>Note that in deriving Eq. (4.1.47), we neglect the effects of multiple internal reflections between the plate surfaces. In practice, most compensators are coated so as to reduce the surface reflection loss. For cases where the direction of the incident wave makes a small angle with the normal to the surface of the retarder and the plane of incidence makes an angle  $\varphi$  with the optical axis of the crystal, Holmes [18] has shown that the phase retardation is given by

$$\delta \cong \frac{2\pi v_0 L \Delta n}{c} \left( 1 - \frac{i^2}{2n_o} \left( \frac{\cos^2(\varphi)}{n_o} - \frac{\sin^2(\varphi)}{n_r} \right) \right)$$



**Figure 4.1.2.** A quarter-waveplate with its slow axis horizontal (i.e., along  $e_1$ ). Assume that the incident light is linearly polarized. The 2 component of the wavefield will then lead the 1 component by  $\pi/2$  and the resulting beam will be right circularly polarized. The plane ( $e_1, e_2$ ) coincides with the crystal surfaces.

circularly polarized. Note that the reverse process works with the same transformation. As the reader can see, the quarter-waveplate transforms linearly polarized light into circularly polarized light, that is,

$$\mathbf{C}\left(\frac{\pi}{2}\right) \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

as illustrated in Fig. 4.1.2. Reversing the direction of propagation of light, a quarter-waveplate will transform circularly polarized light into linearly polarized light. The perspective view in Fig. 4.1.2 identifies the coordinate system and direction of the propagation of the light beam. Two remarks are of importance. First, a quarter-waveplate rotated at frequency  $w$  constitutes a polarization modulator. Second, as the reader can discern, a circular polarizer may be constructed from a linear polarizer followed by a quarter-wave linear compensator with a relative orientation of  $45^\circ$ .

Another example of particular experimental interest is the half-waveplate. The Jones matrix is given by

$$\mathbf{C}(\pi) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (4.1.49)$$

where a  $\pi/2$  phase factor has been ignored. Note the physical meaning of the matrix  $\sigma_1$  as a half-plate:

$$\mathbf{C}(\pi) = i\sigma_1 \quad (4.1.50)$$

A half-waveplate transforms a right-handed circularly polarized wave into a left-handed polarized wave:

$$\mathbf{C}(\pi) \begin{bmatrix} 1 \\ \pm i \end{bmatrix} = \begin{bmatrix} 1 \\ \mp i \end{bmatrix} \quad (4.1.51)$$

Equation (4.1.47) shows that these two optical devices may be constructed by employing birefringent media of appropriate thicknesses.

Now if the principal axes of the retardation plate make an angle  $\theta$  with the  $e_1$  axis, we find, by making use of the rotation matrix, Eq. (4.1.41), that the Jones matrix can now be written, in the rotated frame, carrying out the appropriate similarity transformation, as

$$\mathbf{C}(\delta, \theta) = \mathbf{R}(-\theta)\mathbf{C}(\delta)\mathbf{R}(\theta)$$

$$= \begin{bmatrix} \cos\left(\frac{\delta}{2}\right) + i \sin\left(\frac{\delta}{2}\right) \cos(2\theta) & i \sin\left(\frac{\delta}{2}\right) \sin(2\theta) \\ i \sin\left(\frac{\delta}{2}\right) \sin(2\theta) & \cos\left(\frac{\delta}{2}\right) - i \sin\left(\frac{\delta}{2}\right) \cos(2\theta) \end{bmatrix} \quad (4.1.52)$$

As an example application of Eq. (4.1.52), we consider a half-waveplate rotating at angular velocity  $w$ . For this purpose we set  $\theta = wt$ . Thus Eq. (4.1.52) yields

$$\mathbf{C}(\pi, wt) = i \begin{bmatrix} \cos(2wt) & \sin(2wt) \\ \sin(2wt) & -\cos(2wt) \end{bmatrix} \quad (4.1.53)$$

The effect of a rotating half-waveplate on a circularly polarized wave is now given by

$$\mathbf{C}(\pi, wt) \begin{bmatrix} 1 \\ \pm i \end{bmatrix} = \exp(\pm 2iwt) \begin{bmatrix} 1 \\ \mp i \end{bmatrix} \quad (4.1.54)$$

which should be compared with Eq. (4.1.51). The usefulness of this expression is that the right-hand side has a net time dependence of  $\exp(-i(v_0 \pm 2iw)t)$ , that is, it yields the shifted frequency  $v_0 \pm 2iw$ . The frequency of an incident left-handed circularly wave whose electric field is rotating with the same sense as the wave plate is downshifted by  $2w$ , while the wave of opposite sense is upshifted by  $2w$ .

Many materials are absorbing. We define the Jones matrix of an ideal absorber as

$$\mathbf{A}(\varepsilon, \eta) = \exp(-\eta) \begin{bmatrix} \exp(-\varepsilon) & 0 \\ 0 & \exp(\varepsilon) \end{bmatrix} \quad (4.1.55)$$

with  $\eta \equiv \frac{1}{2}(\eta_1 + \eta_2)$  and  $\varepsilon \equiv \frac{1}{2}(\eta_1 - \eta_2)$ , where  $\eta_1$  and  $\eta_2$  are the absorption coefficients in the 1 and the 2 directions, respectively. The exponential factor  $\exp(-\eta)$  reduces both components of the electric field at the same rate and does not affect the degree of polarization. The matrix  $\sigma_0$  represents an ideal isotropic nonabsorbing optical medium. On substituting Eq. (4.1.55) into Eq. (4.1.16), we obtain the following expression for the gain of an absorber:

$$g = \exp(-2\eta)[\cosh(2\varepsilon) - \langle \sigma_1 \rangle \sinh(2\varepsilon)] \quad (4.1.56)$$

In closing this subsection we pass to a consideration of the Jones matrix of a perfect lossless mirror perpendicular to the beam

$$\mathbf{M} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \sigma_1 \quad (4.1.57)$$

It is important to appreciate that when an incident linear horizontally (resp. vertically) polarized wave is reflected by a perfect mirror, the resultant wave remains linear horizontally (resp. vertically) polarized while the helicity of the emerging light for an incident circularly polarized wave is reversed; that is, the mirror converts a right-handed circular wave into a left-handed circular wave and vice versa.

#### 4.1.1.5. Physical Realizability Constraint

We shall now derive conditions for the physical realizability of passive deterministic Jones matrices. By *physically realizable*, we mean that  $0 \leq g \leq 1$  and  $0 \leq P_o \leq 1$  for all input states of polarization. The central question of concern here is: "What are the physical realizability constraints imposed on the Jones matrix?"

The Hermitian and nonnegative definite characters of  $\mathbf{J}\mathbf{J}^+$  imply a further restriction on its eigenvalues, namely, that they be real and nonnegative. Recall that the eigenvalues of  $\mathbf{J}\mathbf{J}^+$  are identified with the maximum and minimum gain given by Eq. (4.1.10). The corresponding eigenvectors are orthogonal and characterize maximum and minimum gain states of polarization. We will order them by decreasing magnitude such that

$$0 \leq \lambda_a \leq \lambda_b \quad (4.1.58)$$

where

$$\lambda_{a,b} = \frac{1}{2}(\text{tr}(\mathbf{J}^+\mathbf{J}) \pm ((\text{tr}(\mathbf{J}^+\mathbf{J}))^2 - 4|\det(\mathbf{J})|^2)^{1/2}) \quad (4.1.59)$$

As was expressed by Eq. (4.1.20), an alternative expression of the gain in terms

of these eigenvalues is given by

$$g = \frac{\lambda_a + \lambda_b}{2} \quad (4.1.60)$$

The gain for a passive linear medium being less than unity implies the following constraint:

$$0 \leq \lambda_a + \lambda_b \leq 2 \quad (4.1.61)$$

Now, if we make use of the Cauchy–Schwarz inequality, which states that for any two arbitrary matrices  $\mathbf{A}$  and  $\mathbf{B}$  [23]

$$(\text{tr}(\mathbf{AB}))^2 \leq \text{tr}(\mathbf{A}^+ \mathbf{A}) \text{tr}(\mathbf{B}^+ \mathbf{B}) = |\mathbf{A}|^2 |\mathbf{B}|^2 \quad (4.1.62)$$

we arrive at the relation

$$g \leq |\mathbf{J}^+ \mathbf{J}| (\text{tr}(\mathbf{D}_{2i}^2))^2 \leq |\mathbf{J}^+ \mathbf{J}| \quad (4.1.63)$$

where the “Frobenius norm” of  $\mathbf{X}$  is defined by  $|\mathbf{X}| = (\text{tr}(\mathbf{X}^+ \mathbf{X}))^{1/2}$ . Note that the Frobenius norm is neatly characterized in terms of the singular value decomposition (SVD), specifically,  $|\mathbf{A}|^2 = \sum_j \sigma_j^2$ , where  $\sigma_j$  are the singular values of  $\mathbf{A}$ . To make this point more precise, see Appendix F. Now, on substituting the appropriate definition Eq. (3.1.147), we can prove that  $\text{tr}(\mathbf{D}_{2i}^2) = (1 + P^2)/2 \leq 1$ . We see from (F.3) that

$$g \leq (\lambda_a^2 + \lambda_b^2)^{1/2} \quad (4.1.64)$$

Combining Eqs. (4.1.63) and (4.1.64), we may readily express the set of conditions (4.1.61) and (4.1.63) explicitly as

$$\det(\mathbf{J}^+ \mathbf{J}) = |\det(\mathbf{J})|^2 \leq 1 \quad (4.1.65)$$

The usefulness of this expression is that the left-hand side can be evaluated in terms of the eigenvalues. Thus identifying a physically realizable Jones matrix can be accomplished by calculating its eigenvalues and then testing to see whether the two eigenvalues satisfy Eqs. (4.1.61) and (4.1.65).

#### 4.1.1.6. Polar Decomposition of Jones Matrices

Like any square matrix with complex elements, a Jones matrix can be represented in each of the two following ways [18]:

$$\mathbf{J} \equiv \mathbf{H}\mathbf{U} \quad (4.1.66a)$$

or

$$\mathbf{J} \equiv \mathbf{VL} \quad (4.1.66b)$$

where  $\mathbf{H}$  ( $\mathbf{H}^2 = \mathbf{JJ}^+$ ) and  $\mathbf{L}$  ( $\mathbf{L}^2 = \mathbf{J}^+\mathbf{J}$ ) are Hermitian matrices with non-negative spectra, and  $\mathbf{U}$  and  $\mathbf{V}$  are unitary matrices. Equations (4.1.66a, b) are the matrix analogs of the representation of a complex number in the form  $z = |z| \exp(i \arg(z))$ .<sup>7</sup> The matrices  $\mathbf{H}$  and  $\mathbf{L}$  are uniquely determined by the Jones matrix:  $\mathbf{H}$  is the left modulus of  $\mathbf{J}$  while  $\mathbf{L}$  is the right modulus of  $\mathbf{J}$ . However,  $\mathbf{U}$  and  $\mathbf{V}$  are not uniquely determined for a singular matrix. If  $\mathbf{J}$  is nonsingular, then  $\mathbf{U}$  and  $\mathbf{V}$  are also uniquely determined and the matrices  $\mathbf{H}$  and  $\mathbf{L}$  will be matrices with positive spectra.

$$\mathbf{U} = \mathbf{H}^{-1}\mathbf{IV} = \mathbf{IL}^{-1} \quad (4.1.67)$$

The physical interpretation of Eqs. (4.1.66) is that an arbitrary polarization element represented by a Jones matrix can be interpreted as a series of compensator (unitary Jones matrix) followed by a transformation that performs contraction or dilatations in the direction of the eigenvectors of  $\mathbf{H}$  or  $\mathbf{L}$ . In other words, the polar decomposition of a Jones matrix indicates that any optical element is optically equivalent to a series of a linear polarizer  $\mathbf{P}(\alpha)$  and a compensator  $\mathbf{C}(\delta)$  in any of the two possible relative orders. The matrix multiplications, given by Eqs. (4.1.66), are not commutative, and thus we note that the polar decomposition is order-dependent. Since the gain and the output degree of polarization of a unitary Jones matrix are unaltered by varying the incident polarization state,  $\mathbf{U}$  characterizes the transformation of the state of polarization, while  $\mathbf{H}$  characterizes either the polarization ( $P_o > P_i$ ) or the depolarization ( $P_o < P_i$ ).

Now we return to the definition of the diattenuation to remark that [12]

$$d(\mathbf{J}) = d(\mathbf{H}) = d(\mathbf{L}) = \frac{g_{\max} - g_{\min}}{g_{\max} + g_{\min}} = \left(1 - \frac{4|\det(\mathbf{J})|^2}{(\text{tr}(\mathbf{J}^+\mathbf{J}))^2}\right)^{1/2} \quad (4.1.68)$$

Similarly, one can express the retardation as

$$\begin{aligned} r(\mathbf{J}) = r(\mathbf{U}) = r(\mathbf{V}) &= 2 \cos^{-1} |\frac{1}{2} \text{tr}(\mathbf{U})| = 2 \cos^{-1} |\frac{1}{2} \text{tr}(\mathbf{V})| \\ &= 2 \cos^{-1} \left( \frac{\left| \text{tr}(\mathbf{J}) + \frac{\det(\mathbf{J})}{|\det(\mathbf{J})|} \text{tr}(\mathbf{J}^+) \right|}{2(\text{tr}(\mathbf{J}^+\mathbf{J}) + 2|\det(\mathbf{J})|)^{1/2}} \right) \end{aligned} \quad (4.1.69)$$

Reference is made to Whitney for the details of an illustrating example of the polar decomposition to see how  $\mathbf{U}$  and  $\mathbf{H}$  can be operationally obtainable [19].

<sup>7</sup>However, a noticeable difference between matrix and complex number representations is that matrices generally do not commute.

For the sake of completeness, we indicate several theorems about the decomposability of arbitrary Jones matrices into sequences of linear polarizers, rotators, and compensators that can be derived from polar decomposition. We leave to the reader the task of proving a few general theorems relating to optical systems of the type under consideration: (1) any general polarizer can be realized (nonuniquely) from a linear polarizer and two compensators, (2) any compensator can be made (nonuniquely) from two compensators, and (3) any phase shifter can be made from a rotator followed by a compensator [3, 19].

#### 4.1.1.7. Optical Propagation in Multilayered Media

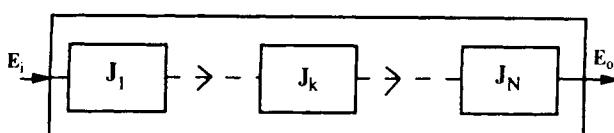
The study of electromagnetic wave propagation in a spatially inhomogeneous medium is of fundamental importance for a wide variety of interesting and potentially useful applications, such as magneto-optical recording [20]. Under certain circumstances inhomogeneous media can be modeled by a series of polarizing elements. Here we consider some general properties of polarization effects in layered optical media with normal incidence. To begin with, let us assume forward-propagating plane waves, that is, reflections, that may give rise to backward-propagating waves to be absent. Note that the assumption of no multiple reflections is legitimate for most practical birefringent networks, in which the crystal plates are relatively thick and the Fresnel reflections are small.

One important property of the Jones matrix formalism is composition; the output of  $\mathbf{J}_k$  can be used as an input of  $\mathbf{J}_{k+1}$ . For  $N$  optical devices in cascade (see Fig. 4.1.3), the transformation of the polarization due to the serial device can be expressed as a single Jones matrix that is simply the product of the  $N$  ordered Jones matrices of the individual optical elements

$$\mathbf{J} = \prod_{k=N}^1 \mathbf{J}_k \quad (4.1.70)$$

Consequently the evolution of the wave is handled by a succession of matrix multiplications. A word of caution is in order here. The decomposition of a Jones matrix into various matrices is not without interpretation problems since, in general, this decomposition is non-commutative.

Let us proceed to examine the consequences of Eq. (4.1.70). For this purpose, we consider the special case  $\mathbf{J}_k = \mathbf{J}_1$  for  $k \in \langle 1, N \rangle$ . Then  $\mathbf{J} = \mathbf{J}_1^N$ . The



**Figure 4.1.3.** Forward-propagating plane wave in a cascade of  $N$  optical systems.

problem reduces to raising  $J_1$  to power  $N$ . To solve this problem, we can use the Jordan canonical form. Suppose that  $J_1$  is transformed into Jordan canonical form  $P$  by  $Q$ , namely,  $P = Q^{-1}J_1Q$  (or equivalently  $J_1 = QPQ^{-1}$ ). Then we have  $J_1^2 = QP^2Q^{-1}$  and by induction  $J_1^N = QP^NQ^{-1}$ . Let us assume that  $J_1$  is semisimple; then we can write  $P$  as a diagonal matrix

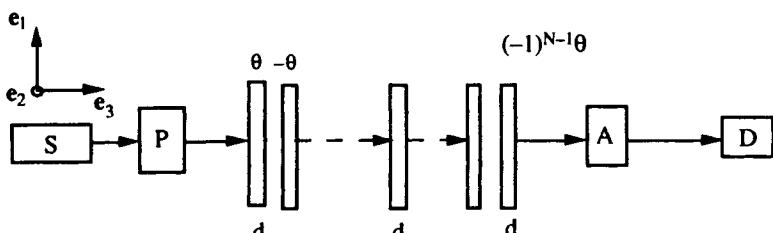
$$P \equiv \begin{bmatrix} p_1 & 0 \\ 0 & p_2 \end{bmatrix}$$

and evidently

$$P^N \equiv \begin{bmatrix} p_1^N & 0 \\ 0 & p_2^N \end{bmatrix}$$

If  $J_1$  is not semisimple, the results are more complicated; reference is made to Pease for the actual expression of  $J_1^N$  [21].

With these remarks, we now outline an evaluation of the gain for a birefringent multilayered medium. For that purpose, we consider a stack of  $N$  identical birefringent retardation plates with two crystal axes oriented in the plane ( $e_1, e_2$ ), with the whole stack located between two crossed polarizers. This geometric arrangement, known as the type I Solc filter (also termed folded Solc filter), is sketched in Fig. 4.1.4 [11, 22]. The orientation angles of the different components are such that  $\theta_k = (-1)^{k-1}\theta$  with  $(k \in \langle 1, N \rangle)$ . In making this calculation we assume that (1) optical interface effects are neglected and (2) multiple reflections between dielectric plane interfaces are neglected. Observe also that the use of the Jones formalism requires that the thickness of each layer  $d_k = d$  be much larger than the wavelength of the illuminating light.



**Figure 4.1.4.** Schematic illustration of the notations relating to the type I Solc filter  $\theta_k = (-1)^{k-1}\theta$   $k \in \langle 1, N \rangle$ ,  $d_k = d \forall k$  with angle oriented with respect  $e_1$ . The multilayered medium is placed between crossed polarizers:  $P(\alpha = 0)$  and  $A = P(\alpha = \pi/2)$ .

We first evaluate the overall Jones matrix for this serial device assuming that  $N$  is an even number. This leads to

$$\mathbf{J} = \mathbf{P} \left( \frac{\pi}{2} \right) (\mathbf{C}(\delta, -\theta) \mathbf{C}(\delta, \theta))^{N/2} \mathbf{P}(0) = \mathbf{P} \left( \frac{\pi}{2} \right) \mathbf{C} \mathbf{P}(0) \quad (4.1.71)$$

where  $\delta$  is the phase retardation of each component. If we substitute Eq. (4.1.52) into the definition of the unimodular matrix  $\mathbf{C}$  [i.e.,  $\det(\mathbf{C}) = 1$ ], in Eq. (4.1.71), we obtain

$$\mathbf{C} = \begin{bmatrix} a & b \\ -b & a^* \end{bmatrix}^{N/2} \quad (4.1.72)$$

where we have set

$$a = \left( \cos\left(\frac{\delta}{2}\right) + i \sin\left(\frac{\delta}{2}\right) \cos(2\theta) \right)^2 + \sin^2(2\theta) \sin^2\left(\frac{\delta}{2}\right) \quad (4.1.73a)$$

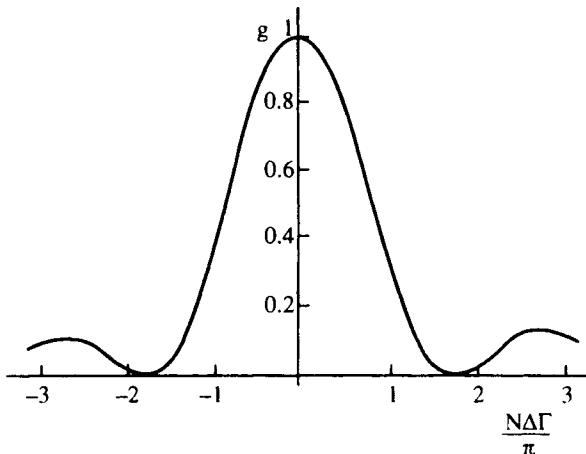
$$b = -\sin^2\left(\frac{\delta}{2}\right) \sin(4\theta) \quad (4.1.73b)$$

Computation of the  $N$ th power of a  $2 \times 2$  unimodular matrix can be done by using the Chebyshev's identity [23]. Written out more explicitly, Eq. (4.1.72) gives, under the preceding conditions

$$\mathbf{C} = \begin{bmatrix} a \frac{\sin\left(\frac{N}{2}X\right) - \sin\left(\left(\frac{N}{2}-1\right)X\right)}{\sin(X)} & b \frac{\sin\left(\frac{N}{2}X\right)}{\sin(X)} \\ -b \frac{\sin\left(\frac{N}{2}X\right)}{\sin(X)} & a^* \frac{\sin\left(\frac{N}{2}X\right) - \sin\left(\left(\frac{N}{2}-1\right)X\right)}{\sin(X)} \end{bmatrix} \quad (4.1.74)$$

where we have defined  $X \equiv \cos^{-1}(\text{Re}(a))$ .

Next the gain of the Solc filter can be evaluated from Eq. (4.1.14), making use of Eq. (4.1.74). In the limit that  $N$  is much larger than 1 and assuming that  $\theta = \pi/4N$ , one can derive approximate and useful results. In this limit, we



**Figure 4.1.5.** Gain of the Solc filter as computed from Eq. (4.1.75).

arrive at the relation

$$g_N = \left[ \frac{\sin\left(\frac{\pi}{2} \left( \left( 1 + \left( \frac{\Delta\Gamma N}{\pi} \right)^2 \right)^{1/2} \right) \right)}{\left( 1 + \left( \frac{\Delta\Gamma N}{\pi} \right)^2 \right)^{1/2}} \right]^2 \quad (4.1.75)$$

where  $\Delta\Gamma$  is defined as  $\Delta\Gamma \equiv (2\pi v_0/c)(n_e - n_o)d - (2m + 1)\pi$ , assuming that each component is characterized by refractive indices  $n_e$  and  $n_o$  [11]. A graph showing the behavior of the gain is displayed in Fig. 4.1.5 as a function of the dimensionless parameter  $\Delta\Gamma N/\pi$ . The reader is referred to Yariv and Yeh for general properties of the layered-optical media [11].

#### 4.1.1.8. Polarization Ratio Transformations

From what was discussed in Section 3.1.4, we know that the physical information on the ellipse of polarization is contained in the complex polarization ratio. Nothing said up to now has exploited any properties of this quantity to describe the evolution of pure states propagating in nondepolarizing optical systems. The orthogonal basis states of polarization ( $e_u, e_v$ ) that are used to define the Jones matrix and the complex polarization ratio  $Z_{uv}$  may be considered arbitrary. However, for simplicity we will assume the orthogonal 1 and 2 linear polarizations to be the basis states. Consider Eq. (3.1.50), and define the complex variables  $Z_i \equiv E_{i2}/E_{i1}$ ,  $Z_o \equiv E_{o2}/E_{o1}$  (we suppress the label 12 for typographic convenience) characterizing the ellipses of polarization at the input and output of the system. Substituting Eqs. (4.1.1–2), one may prove

that these complex variables are interrelated by the conformal bilinear transformation [24]

$$Z_o = \frac{J_{22}Z_i + J_{21}}{J_{12}Z_i + J_{11}} = \operatorname{Re}(Z_o) + i\operatorname{Im}(Z_o) \quad (4.1.76)$$

This expression is important because it shows that the ellipse of polarization at the output of the optical system is completely determined by the ellipse of polarization at the input and the four complex components of the Jones matrix. The reverse problem of determining the ellipse of polarization at the input of the optical system, knowing the ellipse of polarization at the output, is easily solved, and we obtain the bilinear transformation

$$Z_i = \frac{J_{11}Z_o - J_{21}}{-J_{12}Z_o + J_{22}} \quad (4.1.77)$$

If the transformation is intensity-invariant (i.e.,  $J^+ = J^{-1}$ ), and making use of Eq. (4.1.77), we readily find that

$$Z_o = \frac{J_{11}^*Z_i - J_{12}^*}{J_{12}Z_i + J_{11}} \quad \text{with} \quad |J_{11}|^2 + |J_{12}|^2 = 1 \quad (4.1.78)$$

Our interest is now in the polarization orthogonalization properties of optical systems, that is, to find the conditions under which a nondepolarizing optical system described by a Jones matrix may transform an input state of polarization into an output state of polarization that is orthogonal. The analysis is largely based on a paper by Azzam [25]. Following the formalism developed in Section 3.1.4, we know that the states  $Z_i$  and  $Z_o$  are orthogonal if

$$Z_o Z_i^* = -1 \quad (4.1.79)$$

On substitution from Eq. (4.1.76) into Eq. (4.1.79), we have immediately

$$J_{22}|Z_i|^2 + J_{12}Z_i + J_{21}Z_i^* + J_{11} = 0 \quad (4.1.80)$$

Now we make the following substitutions

$$J_{kl} = r_{kl} + is_{kl} \quad (4.1.81)$$

$$Z_i = x + iy \quad (4.1.82)$$

On separating Eq. (4.1.80) into its real and imaginary parts, we get a pair of equations

$$r_{22}(x^2 + y^2) + (r_{12} + r_{21})x + (s_{21} - s_{12})y + r_{11} = 0 \quad (4.1.83a)$$

$$s_{22}(x^2 + y^2) + (r_{12} - r_{21})y + (s_{12} + s_{21})x + s_{11} = 0 \quad (4.1.83b)$$

that should be satisfied simultaneously. As a specific illustration of these formulas, we consider the Jones matrix of the  $\pi/2$  rotator, Eq. (4.1.44),

$$\mathbf{J} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

Substituting this matrix into Eqs. (4.183a, b), we have  $y = 0$ . Consequently, a  $\pi/2$  rotator orthogonalizes all input states represented by all points on the real axis of the  $Z_i$ -plane, that is, all input linear polarization states. The reader is referred to Azzam's paper [25] for a detailed discussion of the general polarization orthogonalization properties of optical systems.

#### 4.1.1.9. Jones Calculus and Entropy Transformations

We now pass on to the problem of evaluating the variation of polarization entropy, introduced by the interaction of the wave with any linear optical system that can be described by a Jones matrix, in closed form. For that purpose, we substitute Eq. (4.1.11) into Eq. (3.4.5). The result is

$$\Delta S = S_o - S_i = \ln \left( \frac{s(P_i)}{s(P_o)} \right) = \ln \left( \frac{s(P_i)}{s \left( (1 - (1 - P_i^2) \frac{|\det(\mathbf{J})|^2}{g^2})^{1/2} \right)} \right) \quad (4.1.84)$$

which indicates that the variation of entropy is governed by both the degree of polarization of the input light and the characteristics of the optical medium. It should be further noted that entropy conservation during the interaction (i.e.,  $\Delta S = 0$ ) implies invariance of the degree of polarization and conversely. If we consider a unitary Jones medium ( $|\det(\mathbf{J})| = g = 1$ ), then from the property of similarity invariance (invariance of the trace), we see that it is also entropy-invariant. We arrive to the result that a depolarizing interaction, namely,  $P_o < P_i \neq 1$ , induces a loss of polarization entropy [26].

Our task is now to consider the problem of characterizing the polarizance of an optical medium, specifically, the coupling of unpolarized light into polarized light. For an unpolarized incident state, the variation of entropy is expressed as

$$\Delta S_u = -\ln(2) - \ln \left( s \left( \left( 1 - \frac{4|\det(\mathbf{J})|^2}{(\text{tr}(\mathbf{J}^+ \mathbf{J}))^2} \right)^{1/2} \right) \right) = -\ln(2s(d)) \quad (4.1.85)$$

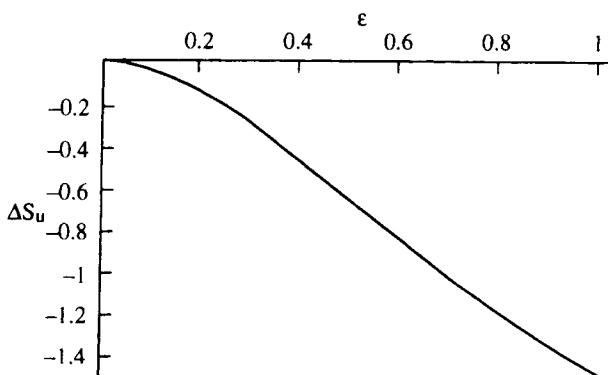
where  $d$  denotes the diattenuation.

We now consider briefly a few familiar examples to illustrate the implications of Eq. (4.1.85) by applying it to the basic optical components: the compensator  $\mathbf{C}(\delta)$ , the rotator  $\mathbf{R}(\theta)$ , the linear polarizer  $\mathbf{P}(x)$ , and the absorber  $\mathbf{A}(\varepsilon, \eta)$ . As we have just remarked, the moduli of  $\lambda_a$  and  $\lambda_b$  are equal for both the compensator and the retarder (see Table 4.1.1). Consequently we have  $\Delta S_u = 0$ . It turns out that for a polarizer,  $\Delta S_u$  is constant and fixed to  $-\ln(2)$ . This is an especially satisfying result, as it is precisely what is required by the two-level description of partially polarized light. In view of Eq. (4.1.85), one may write the variation of polarization entropy, for an incident unpolarized wave on an absorber as

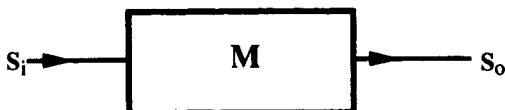
$$\begin{aligned}\Delta S_u &= -\ln(2) - \ln\left(s\left(\left(1 - \frac{1}{(\cosh(2\varepsilon))^2}\right)^{1/2}\right)\right) \\ &= \ln[\cosh(2\varepsilon)] - 2\varepsilon \tanh(2\varepsilon)\end{aligned}\quad (4.1.86)$$

The behavior of  $\Delta S_u(\varepsilon)$  is shown in Fig. 4.1.6. The curve undergoes a monotone decrease. We get  $\Delta S_u = 0$  when  $\varepsilon = 0$ , corresponding to the isotropic case whose gain reads  $g_u = \text{tr}[\mathbf{A}(\varepsilon, \eta)\mathbf{A}^+(\varepsilon, \eta)] = \exp(-2\eta)$ .

In summary, we emphasize the fact that depolarization is connected to a process of *entropy production*. The preceding analysis has shown that depolarization is connected to a relation of partial ordering in the set of density matrices. A decrease of the purity character [i.e.,  $\mathbf{D}_{2i}(P_i) \rightarrow \mathbf{D}_{2o}(P_o < P_i)$ ] implies an increase of the entropy and a contracting subspace. On the other hand, an increase of the purity character [i.e.,  $\mathbf{D}_{2i}(P_i) \rightarrow \mathbf{D}_{2o}(P_o > P_i)$ ] leads to a dilating subspace. Within this geometric interpretation, it is important to appreciate that pure states are invariant states of the dilating case, whereas the unique invariant of the contracting interaction is the totally unpolarized state where everything is concentrated at the origin. Note that it may be an objective of optical design in polarimetry to build a measurement system with  $\Delta S \cong 0$ .



**Figure 4.1.6.** Variation of polarization entropy,  $\Delta S_u$ , for an incident unpolarized wave on an absorber as a function of  $\varepsilon$ .



**Figure 4.1.7.** Black box representing an incident light beam, of Stokes vector  $S_i$ , interacting with an optical medium, characterized by a Mueller matrix  $M$ , to give an emerging beam  $S_o$ .

#### 4.1.2. MUELLER CALCULUS

The effect of an optical device on the polarization of light can be alternately characterized by a Mueller matrix. The description of optical systems in terms of Mueller matrices is applicable to more general situations than is the description in terms of Jones matrices. We can anticipate that some optical devices can be described by Mueller matrices but not by Jones matrices, and a formal proof will be given shortly. The subclass of Mueller matrices that can be derived from Jones matrices is termed the *Mueller-Jones set*. This observation will be important for interpretation later on.

When a suitable coordinate system is established, the optical medium can be mathematically represented by a  $4 \times 4$  Mueller matrix that acts on the incident Stokes vector (see also Fig. 4.1.7).

Let the input and the output states of polarization be parametrized by the Stokes vector  $S_i$  and  $S_o$ . We assume that the matrix  $M$  acts on the input state  $S_i$  by matrix multiplication to give the output state

$$S_o = MS_i \quad (4.1.87)$$

where the  $4 \times 4$  real-valued  $M$  matrix

$$M = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix} \quad (4.1.88)$$

termed the *Mueller matrix of the optical element*,<sup>8</sup> is a characteristic of the optical medium. In general, it depends on frequency and in case of light scattering exhibits an angular dependence. In general, all 16 matrix elements  $m_{jk}$  are independent; however, imposition of symmetry conditions on  $M$ , or depending on certain optical properties of  $M$ , will reduce the number of independent parameters (some  $m_{kj}$  could be equal to others, and some might be zero). The output Stokes parameters are linear combinations of the

<sup>8</sup>However, Soleillet was historically the first to consider the properties of  $M$  matrices [27].

input Stokes parameters, each weighted by the corresponding Mueller matrix elements. The element  $m_{kj}$ , in the  $k$ th row and  $j$ th column, describes how much of the  $j$ th input Stokes parameter is converted to the  $k$ th output Stokes parameter. For instance,  $m_{34}$  describes how the Mueller matrix  $\mathbf{M}$  acts like a quater-wave compensator. Bear in mind that not every  $4 \times 4$  real matrix corresponds to a Mueller matrix. There are several constraints on the  $m_{kj}$  elements that are very useful as a check on the experimental determinations or numerical computations. For instance, these elements obey the rule

$$|m_{kj}| \leq m_{00} \leq 1 \quad (4.1.89)$$

We will explain more about that in a moment.

The overall Mueller matrix  $\mathbf{M}$  for a cascade of  $N$  systems, each having a Mueller matrix  $\mathbf{M}_k$ ,  $k \in \langle 1, N \rangle$  can be expressed as the reverse order matrix product:

$$\mathbf{M} = \prod_{k=N}^1 \mathbf{M}_k \quad (4.1.90)$$

The methods employed for measuring the different elements of the Mueller matrix will be described in Chapter 5.

#### 4.1.2.1. Mueller Formalism and $4 \times 4$ Coherency Matrix Formulation

In this section we describe the procedure for calculating the change of the  $4 \times 4$  coherency matrix when a quasimonochromatic wave is transmitted through a Mueller medium. Almost all experiments designed to extract polarization information make use of second-order statistics, but they may be insufficient for many purposes, such as intensity correlations experiments. A natural descriptor of the fluctuations in light is the  $4 \times 4$  coherency matrix. In Part 3, we have obtained the fourfold equivalent of the  $2 \times 2$  coherency matrix formalism. This  $4 \times 4$  Hermitian matrix contains information on variance and correlation of the instantaneous Stokes parameters. We now consider the problem of the relation of the  $4 \times 4$  coherency matrix to the Mueller matrix formalism [28]. Taking the direct product of each side of Eq. (4.1.87) times its Hermitian adjoint followed by ensemble averaging yields

$$\Phi_{4o} = \langle \mathbf{S}_o \otimes \mathbf{S}_o^+ \rangle - \langle \mathbf{S}_o \rangle \otimes \langle \mathbf{S}_o^+ \rangle = \mathbf{M} \Phi_{4i} \mathbf{M}^T \quad (4.1.91)$$

This equation provides the basic law of interaction of a partially polarized wavefield with a deterministic linear optical device. This congruent transformation constitutes a generalization of Eq. (4.1.7). Let us consider the extremes in polarization behavior. If the incident Gaussian wavefield is unpolarized, we

obtain

$$\Phi_{4\text{ogu}} = \frac{\langle S_0 \rangle_u^2}{4} \mathbf{M} \mathbf{M}^T \quad (4.1.92)$$

On the other hand, for a completely polarized wavefield, we find

$$\det(\Phi_{4\text{ogp}}) = 0 \quad (4.1.93)$$

which parallels Eq. (4.1.12). An interesting problem would be the extension of the analysis to cover the case of non-Gaussian wavefields where one would have to deal with Eq. (4.1.91) without the simplification afforded by the Gaussian assumption. We shall not develop these aspects further as the corresponding expressions for the  $4 \times 4$  coherency matrix in the situation of non-Gaussian statistics are rather involved.

#### 4.1.2.2. Gain

In view of our discussion of the gain within the Jones formalism, we can ask for the expression of this quantity when the optical system is characterized by a Mueller matrix. The first row of Eq. (4.1.41) yields

$$g = m_{00} + \sum_{k=1}^3 m_{0k} \langle \sigma_k \rangle \quad (4.1.94)$$

in terms of the  $m_{0k}$  and the input normalized Stokes parameters ( $\langle \sigma_k \rangle \equiv \langle S_k \rangle / \langle S_0 \rangle$ ). Observe that the first row of a Mueller matrix determines the gain completely. As an immediate application of Eq. (4.1.94), for an input unpolarized radiation, the gain reduces to

$$g_u = m_{00} \quad (4.1.95)$$

Note that Eq. (4.1.95) yields a passive system requirement in the form of an inequality

$$0 \leq m_{00} \leq 1 \quad (4.1.96)$$

However, this inequality cannot be tested in practice because of the unfortunate habit of normalizing  $m_{00}$  to unity.

By an argument similar to that used in deriving Eqs. (4.1.18a, b), the

maximum and minimum gain occur when the vectors

$$\begin{bmatrix} m_{01} \\ m_{02} \\ m_{03} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \langle \sigma_1 \rangle \\ \langle \sigma_2 \rangle \\ \langle \sigma_3 \rangle \end{bmatrix}$$

are parallel and antiparallel and can be shown to be given by

$$g_{\max} = m_{00} + P_i \left( \sum_{k=1}^3 m_{0k}^2 \right)^{1/2} \quad (4.1.97a)$$

$$g_{\min} = m_{00} - P_i \left( \sum_{k=1}^3 m_{0k}^2 \right)^{1/2} \quad (4.1.97b)$$

The associated orthogonal Stokes vectors can be easily derived, and we obtain

$$\mathbf{S}_{\max} = \frac{1}{P_i (\sum_{k=1}^3 m_{0k}^2)^{1/2}} \begin{bmatrix} 1 \\ m_{01} \\ m_{02} \\ m_{03} \end{bmatrix} \quad (4.1.98a)$$

$$\mathbf{S}_{\min} = \frac{1}{P_i (\sum_{k=1}^3 m_{0k}^2)^{1/2}} \begin{bmatrix} 1 \\ -m_{01} \\ -m_{02} \\ -m_{03} \end{bmatrix} \quad (4.1.98b)$$

In analogy with Eq. (4.1.25), one obtains from Eqs. (4.1.97a, b)

$$\frac{g_{\max} - g_{\min}}{g_{\max} + g_{\min}} = P_i \frac{(\sum_{k=1}^3 m_{0k}^2)^{1/2}}{m_{00}} \quad (4.1.99)$$

#### 4.1.2.3. Physical Realizability Conditions

Not all  $4 \times 4$  real matrices qualify as Mueller matrices. The Mueller matrix, if properly measured, characterizes the optical medium so that the effect on polarization properties of any beam may be found by means of Eq. (4.1.41). The purpose of this section is to discuss a method for specifying the physical realizability condition of Mueller matrices. Our approach will be phenomenologic in that we assume a complete knowledge of the Mueller matrix from the outset and seek to answer the following question: What are the conditions on a given Mueller matrix that ensure that the output light has a degree of

polarization that does not exceed unity for any polarization of the input light? These constraints are particularly important for calibration of polarimetric devices and testing computational procedures. They can be used to identify erroneous measurement results and to detect computational mistakes.<sup>9</sup> Note that, in many techniques that have been developed for measuring the Mueller polarization matrices, the effects are quite small, and their characterization requires careful calibration and measurement procedure. It is also worth observing that while the diagonal representation of  $\mathbf{J}$  has important physical interpretation, diagonalization of  $\mathbf{M}$  has no such clear physical meaning. In particular, the conditions for physical realizability cannot be viewed as a statement of the nonnegative nature of the eigenvalues of  $\mathbf{M}$ . In fact, it can be shown that it is the eigenvalues of  $\mathbf{GM}^T\mathbf{GM}$  that are important in determining the physical realizability constraints for  $\mathbf{M}$  [29, 32]. A given Mueller matrix  $\mathbf{M}$  is not overpolarizing if and only if the spectrum of  $\mathbf{GM}^T\mathbf{GM}$  is real and an eigenvector associated with the largest eigenvalue is a physical Stokes vector, that is, its degree of polarization does not exceed unity [32].

These constraints are complex to analyze, and, rather than develop a formal development, we chose instead to outline a simple phenomenological approach in that we assume a complete knowledge of the Mueller matrix from the outset. We restrict our analysis to incident Stokes vectors in the form of pure states, which is usually the case encountered in practice for measuring a Mueller matrix. Let the incident pure state be written in terms of the azimuthal and ellipticity angles via Eqs. (3.1.151)–(3.1.153)

$$\mathbf{S}_i = \begin{bmatrix} \langle S_0 \rangle \\ \langle S_0 \rangle \cos(2\chi_i) \cos(2\psi_i) \\ \langle S_0 \rangle \cos(2\chi_i) \sin(2\psi_i) \\ \langle S_0 \rangle \sin(2\chi_i) \end{bmatrix} \quad (4.1.100)$$

and define a function  $F_j(\chi_i, \psi_i)$  by the equation

$$F_j(\chi_i, \psi_i) = m_{j0} + m_{j1} \cos(2\chi_i) \cos(2\psi_i) + m_{j2} \cos(2\chi_i) \sin(2\psi_i) + m_{j3} \sin(2\chi_i) \quad (4.1.101)$$

The gain that follows with the help of Eq. (4.1.101) is derived directly from the relation (4.1.13)

$$g(\chi_i, \psi_i) = F_0(\chi_i, \psi_i) \quad (4.1.102)$$

<sup>9</sup>The search of necessary conditions in the form of inequalities on the elements of a general Mueller matrix was already recognized and discussed in the literature, in connection with the physical realizability condition. However, these conditions are often input-dependent; the fact that the degree of polarization  $P_u = (m_{10}^2 + m_{20}^2 + m_{30}^2)^{1/2}/m_{00}$  produced by an optical system when the incident light is unpolarized is less than unity implies that the following necessary condition must be satisfied by a physically realizable Mueller matrix:  $(m_{10}^2 + m_{20}^2 + m_{30}^2)^{1/2} \leq m_{00} \leq 1$ .

and the output degree of polarization is expressed as

$$P_o(\chi_i, \psi_i) = \frac{1}{F_o(\chi_i, \psi_i)} \left( \sum_{j=1}^3 (F_j(\chi_i, \psi_i))^2 \right)^{1/2} \quad (4.1.103)$$

By definition, the Mueller matrix  $\mathbf{M}$  for the polarization modification device is physically realizable if and only if  $\forall \chi_i$  and  $\psi_i$  such that  $-(\pi/4) \leq \chi_i \leq \pi/4$  and  $0 \leq \psi_i \leq \pi$ , we have

$$0 \leq g(\chi_i, \psi_i) \leq 1 \quad (4.1.104a)$$

$$0 \leq P_o(\chi_i, \psi_i) \leq 1 \quad (4.1.104b)$$

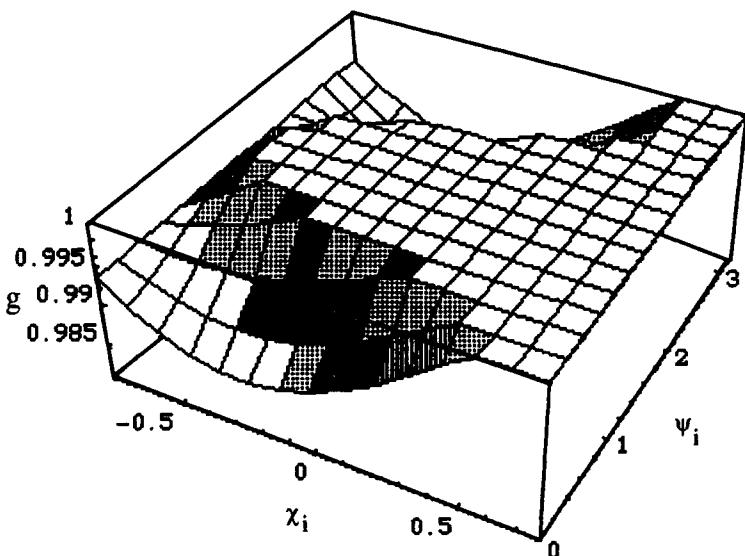
The first constraint, Eq. (4.1.104a), is termed the *overgain condition*, namely, the problem of ensuring that the gain is less than unity and the second constraint, Eq. (4.1.104b), is the *overpolarization condition*, namely, the problem of ensuring that the output degree of polarization does not exceed unity. If one of these constraints is violated, one can safely assume that there are random errors (noise) or/and systematic errors involved in the measurement of the Mueller matrix. In that case it appears as a general rule that any device imperfection can be modeled by postmultiplying the Mueller matrix for the ideal instrument by another Mueller matrix, and this can be used to improve future performance of the instrument.

In order to illustrate the physical realizability conditions, Eqs. (4.1.104a, b), they will be examined in the context of a specific example. For this purpose, consider the following experimentally determined Mueller matrix from a measurement of reflection from a dielectric underwater target [31].

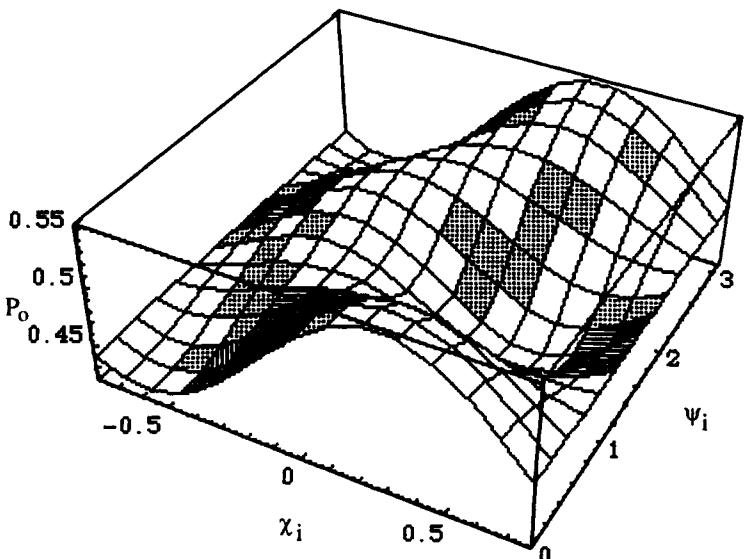
$$\mathbf{M} = \begin{bmatrix} 1 & -0.0118 & 0.0059 & 0.0066 \\ 0.0226 & 0.494 & -0.0054 & 0.0077 \\ 0.0061 & -0.0188 & -0.5055 & -0.007 \\ -0.0154 & -0.006 & 0.0006 & -0.4492 \end{bmatrix} \quad (4.1.105)$$

The rms (root mean square) deviation of the larger diagonal elements of this matrix is  $\pm 0.007$ . Note that all elements of this matrix have been divided by the 00 entry. We have performed calculations of the gain and output degree of polarization by spanning the Stokes space of input states of polarization. The results, respectively given in the three-dimensional plots in Figs. 4.1.8 and 4.1.9, show that this matrix satisfies the passive requirement constraint and does not overpolarize for any incident Stokes vectors.

Thus we conclude that the Mueller matrix, given by Eq. (4.1.105), passes the tests of physical realizability. An alternative approach to the overpolarization problem, via an eigenvalue condition on the  $\mathbf{GM}^T \mathbf{GM}$  matrix, has been



**Figure 4.1.8.** Three-dimensional plot of the gain of the optical medium characterized by the Mueller matrix given by Eq. (4.1.55) as a function of the incident state of polarization parametrized by the angles  $\chi_i$  and  $\psi_i$ .



**Figure 4.1.9.** Output degree of polarization of the optical medium characterized by the Mueller matrix given by Eq. (4.1.55) as a function of the incident state of polarization parametrized by the angles  $\chi_i$  and  $\psi_i$ .

developed by Givens and Kostinski [32]. The numerical results obtained here are in coincidence with the positive verdict on this Mueller matrix by the preceding analysis and condition on the  $\mathbf{GM}^T\mathbf{GM}$  matrix.

#### 4.1.2.4. Mueller Matrices for Common Optical Devices

We again try to derive the Mueller matrix representations of a group of components that modify the state of polarization of incident illumination. We are interested only in elements that are used primarily in the synthesis and analysis of light. We follow a procedure that parallels the Jones matrix case.

We first consider the Mueller matrix of an ideal homogeneous linear compensator. From the definition of the Stokes parameters and by making use of Eq. (4.1.34), we find that

$$\begin{aligned} S_{00} &= |E_{o1}|^2 + |E_{o2}|^2 = |E_{i1}|^2 + |E_{i2}|^2 = S_{i0} \\ S_{01} &= |E_{o1}|^2 - |E_{o2}|^2 = |E_{i1}|^2 - |E_{i2}|^2 = S_{i1} \\ S_{02} &= E_{o1}E_{o2}^* + E_{o2}E_{o1}^* = \cos(\delta)S_{i2} - \sin(\delta)S_{i3} \\ S_{03} &= i(E_{o1}E_{o2}^* - E_{o2}E_{o1}^*) = \sin(\delta)S_{i2} + \cos(\delta)S_{i3} \end{aligned} \quad (4.1.106)$$

where  $\delta$  denotes the differential phase shift  $\delta = \delta_1 - \delta_2$  between the two orthogonally linearly polarized components of the optical field. Thus a compensator has a Mueller matrix given by

$$\mathbf{C}(\delta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(\delta) & -\sin(\delta) \\ 0 & 0 & \sin(\delta) & \cos(\delta) \end{bmatrix} \quad (4.1.107)$$

In view of our discussion of the Jones matrices for common optical devices, we write the Mueller matrix for a compensator with retardance  $\delta$  and with fast axis at an azimuthal angle  $\theta$  of rotation with respect to the horizontal axis:

$$\mathbf{C}(\delta, \theta) =$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2(2\theta) + \sin^2(2\theta)\cos(\delta) & \sin(2\theta)\cos(2\theta)(1 - \cos(\delta)) & -\sin(2\theta)\sin(\delta) \\ 0 & \sin(2\theta)\cos(2\theta)(1 - \cos(\delta)) & \sin^2(2\theta) + \cos^2(2\theta)\cos(\delta) & \cos(2\theta)\sin(\delta) \\ 0 & \sin(2\theta)\sin(\delta) & -\cos(2\theta)\sin(\delta) & \cos(\delta) \end{bmatrix} \quad (4.1.108)$$

As an illustrative example, we consider the Mueller matrix corresponding to a quarter-waveplate with fast axis at an angle of  $0^\circ$ :

$$\mathbf{C}\left(\frac{\pi}{2}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix} \quad (4.1.109)$$

A quarter-waveplate with its fast axis at an angle of  $+45^\circ$  (resp.  $-45^\circ$ ) to a linearly horizontally polarized wave produces a left-hand (resp. right-hand) circularly polarized wave

$$\begin{bmatrix} 1 \\ 0 \\ \pm 1 \\ 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \\ \mp 1 \end{bmatrix} \quad (4.1.110)$$

In like fashion, if the incident wave is right-hand (resp. left-hand) circularly polarized, the emerging wave is linearly  $+45^\circ$  (resp.  $-45^\circ$ ) polarized

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \pm 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 \\ 0 \\ \pm 1 \\ 0 \end{bmatrix} \quad (4.1.111)$$

The Mueller matrix for a half-wave compensator is given by

$$\mathbf{C}(\pi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (4.1.112)$$

As the reader can discern, we have the following transformation of the Stokes vector

$$\mathbf{S}_i = \begin{bmatrix} \langle S_0 \rangle \\ \langle S_1 \rangle \\ \langle S_2 \rangle \\ \langle S_3 \rangle \end{bmatrix} \rightarrow \mathbf{S}_o = \begin{bmatrix} \langle S_0 \rangle \\ \langle S_1 \rangle \\ -\langle S_2 \rangle \\ -\langle S_3 \rangle \end{bmatrix} \quad (4.1.113)$$

Consequently, the output azimuth angle is given by

$$\psi_o = \frac{\pi}{2} - \psi_i \quad (4.1.114a)$$

and the output ellipticity angle is

$$\chi_o = \frac{\pi}{2} + \chi_i \quad (4.1.114b)$$

We next consider a rotator. If we substitute from Eq. (4.1.40) into Eqs. (3.1.112) we then have

$$\begin{aligned} S_{o0} &= |E_{o1}|^2 + |E_{o2}|^2 = |E_{i1}|^2 + |E_{i2}|^2 = S_{i0} \\ S_{o1} &= |E_{o1}|^2 - |E_{o2}|^2 = |E_{i1}|^2 - |E_{i2}|^2 = \cos(2\theta)S_{i1} + \sin(2\theta)S_{i2} \\ S_{o2} &= E_{o1}E_{o2}^* + E_{o2}E_{o1}^* = -\sin(2\theta)S_{i1} + \cos(2\theta)S_{i2} \\ S_{o3} &= i(E_{o1}E_{o2}^* - E_{o2}E_{o1}^*) = S_{i3} \end{aligned} \quad (4.1.115)$$

The Mueller matrix of a rotator can be written in the form

$$\mathbf{R}(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(2\theta) & \sin(2\theta) & 0 \\ 0 & -\sin(2\theta) & \cos(2\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.1.116)$$

As already mentioned, Eq. (4.1.42a), this matrix is unitary, and we have  $\mathbf{R}^{-1}(\theta) = \mathbf{R}(-\theta)$ . A rotation of the Jones vector about the 3-axis through an angle  $\theta$  induces a rotation about the  $\langle S_3 \rangle$ -axis through an angle  $2\theta$  in the Stokes parameters space. It clearly shows that  $\langle S_0 \rangle$ ,  $\langle S_1 \rangle^2 + \langle S_2 \rangle^2$ , and  $\langle S_3 \rangle$  are invariant under this operation. We note that a rotator rotates the polarization ellipse of the incident wave, keeping the ellipticity unchanged. It is simple to prove the following formula for the output azimuth angle

$$\tan(2\psi_o) = \frac{\langle S_{o2} \rangle}{\langle S_{o1} \rangle} = \frac{-\sin(2\theta)\langle S_{i1} \rangle + \cos(2\theta)\langle S_{i2} \rangle}{\cos(2\theta)\langle S_{i1} \rangle + \sin(2\theta)\langle S_{i2} \rangle} = \tan(2\psi_i - 2\theta) \quad (4.1.117)$$

when we have set  $\tan(2\psi_i) = \langle S_{i2} \rangle / \langle S_{i1} \rangle$  according to Eq. (3.1.155c). Thus we obtain

$$\psi_o = \psi_i - \theta \quad (4.1.118)$$

Now we consider the Mueller matrix of an ideal linear polarizer oriented at angle  $\theta$ . From Eqs. (4.1.28) and (3.1.112), we have immediately the form of the Mueller matrix for a linear polarizer

$$\mathbf{P}(\alpha) = \frac{1}{2} \begin{bmatrix} 1 & \cos(2\alpha) & \sin(2\alpha) & 0 \\ \cos(2\alpha) & \cos^2(2\alpha) & \cos(2\alpha)\sin(2\alpha) & 0 \\ \sin(2\alpha) & \cos(2\alpha)\sin(2\alpha) & \sin^2(2\alpha) & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.1.119)$$

As two illustrative examples, consider the cases of an ideal linear horizontal polarizer

$$\mathbf{P}(0) = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.1.120)$$

and of an ideal linear vertical linear polarizer

$$\mathbf{P}\left(\frac{\pi}{2}\right) = \frac{1}{2} \begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.1.121)$$

The gain of the linear polarizer can be evaluated from Eq. (4.1.94). The final result is

$$g = \frac{1}{2}(1 + \cos(2(\alpha - \psi_i)) \cos(2\chi_i)) \quad (4.1.122)$$

where  $\psi_i$  and  $\chi_i$  denote the azimuthal angle and the ellipticity angle of the incident state of polarization, respectively.

We now desire to construct the Mueller matrices of ideal circular polarizers. One common way to derive an expression for the Mueller matrix of a right circular polarizer is to consider that this device is made from a quarter-wave compensator at a relative orientation of  $45^\circ$  and a linear polarizer with its transmission axis  $0^\circ$  in proper combination. The matrix mathematics is

$$\begin{aligned} \mathbf{P}_{cl} &= \mathbf{C}\left(\frac{\pi}{2}, \frac{\pi}{4}\right)\mathbf{P}(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ &= \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix} \end{aligned} \quad (4.1.123)$$

This optical combination produces right-circularly polarized whatever the

incident light beam:

$$\mathbf{S}_i = \begin{bmatrix} \langle S_0 \rangle \\ \langle S_1 \rangle \\ \langle S_2 \rangle \\ \langle S_3 \rangle \end{bmatrix} \rightarrow \mathbf{S}_o = \begin{bmatrix} \langle S_0 \rangle + \langle S_1 \rangle \\ 0 \\ 0 \\ \langle S_0 \rangle + \langle S_1 \rangle \end{bmatrix}$$

However, this matrix is not identical to the form for a “true” circular polarizer

$$\mathbf{P}_{cl} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \quad (4.1.124a)$$

which leads to the transformation

$$\mathbf{S}_i = \begin{bmatrix} \langle S_0 \rangle \\ \langle S_1 \rangle \\ \langle S_2 \rangle \\ \langle S_3 \rangle \end{bmatrix} \rightarrow \mathbf{S}_o = \begin{bmatrix} \langle S_0 \rangle + \langle S_3 \rangle \\ 0 \\ 0 \\ \langle S_0 \rangle + \langle S_3 \rangle \end{bmatrix}$$

It is important to notice that both matrices create a right-circularly polarized Stokes vector, but only Eq. (4.1.124a) brings the information about circular polarization, the third Stokes parameter  $\langle S_3 \rangle$ , creating a total intensity signal  $\langle S_0 \rangle + \langle S_3 \rangle$ . The other one, Eq. (4.1.123), creates the Stokes vector that contains information about horizontal polarization even though it is carried in a circularly polarized beam. In a similar way, the Mueller matrix for a left-circular polarizer can be written as

$$\mathbf{P}_{cr} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \quad (4.1.124b)$$

We next consider the Mueller matrix of a partial linear polarizer:

$$\mathbf{P}(\alpha) = \begin{bmatrix} 1 & \cos(2\alpha) & 0 & 0 \\ \cos(2\alpha) & 1 & 0 & 0 \\ 0 & 0 & \sin(2\alpha) & 0 \\ 0 & 0 & 0 & \sin(2\alpha) \end{bmatrix} \quad (4.1.125)$$

For a linear horizontal polarizer,  $\alpha = 0$  and for a linear vertical polarizer  $\alpha = \pi/2$ . The output degree of polarization of the wave emerging from a partial linear polarizer can be expressed as

$$P_o = \left( \frac{\cos^2(2\alpha)(\langle S_0 \rangle^2 + \langle S_1 \rangle^2) + 2 \cos(2\alpha)\langle S_0 \rangle \langle S_1 \rangle + \sin(2\alpha)P_i^2 \langle S_0 \rangle^2}{\langle S_0 \rangle^2 + \cos^2(2\alpha)\langle S_1 \rangle^2 + 2 \cos(2\alpha)\langle S_0 \rangle \langle S_1 \rangle} \right)^{1/2}$$

Referring to the definition of the ideal absorber, Eq. (4.1.55), we readily find that it can be written in Mueller matrix form as

$$A(\eta, \varepsilon) = \exp(-2\eta) \begin{bmatrix} \cosh(2\varepsilon) & -\sinh(2\varepsilon) & 0 & 0 \\ -\sinh(2\varepsilon) & \cosh(2\varepsilon) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.1.126)$$

We next consider a perfect lossless and nondepolarizing mirror. On substitution from Eq. (4.1.57) into Eq. (4.1.87) and making use of Eq. (3.1.112a-d), we arrive at the relation

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (4.1.127)$$

Finally we outline a derivation of the Mueller matrix of depolarizers. The diagonal matrix given by Eq. (4.1.129) represents an ideal depolarizer that completely depolarizes all incident states of polarization regardless of their form of polarization:

$$\mathbf{D} = \begin{bmatrix} m_{00} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.1.128)$$

Such ideal depolarizers do not exist. However, we anticipate that this polarization matrix can be approximated by multiple scattering from a random particulate medium. Some further commentary on this point is postponed until Section 4. We now pass to a consideration of special depolarizing optical

systems. We can show that

$$\mathbf{D}_L = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix} \quad (4.1.129)$$

yields a circular output Stokes vector for arbitrary incident light; this element may be called a *linear pseudodepolarizer*. Note that this matrix has eight degrees of freedom. In like fashion, the exiting Stokes vector from the optical system having the Mueller matrix

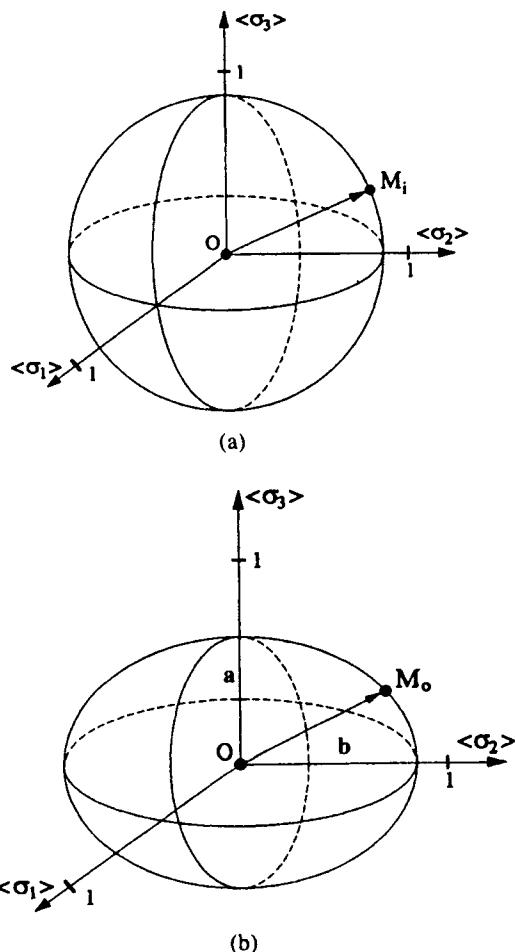
$$\mathbf{D}_C = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.1.130)$$

is purely linear. The matrix given by expression (4.1.130) can be considered as a *circular pseudodepolarizer*. This matrix has twelve degrees of freedom. We caution the reader that these examples of pseudodepolarizers indicate that they may have certain polarizing capabilities [33].

#### 4.1.2.5. Mueller Matrices and Their Parameterizations Within Poincaré Space

In this section we present an alternative approach to the analytic description of deterministic optical systems that offers a geometric point of view. The key concept to this approach is that the Pauli matrices underly the geometry of the Poincaré sphere. For our purpose here, we recall that the normalized Stokes parameters define a three-dimensional vector  $[\langle\sigma_1\rangle, \langle\sigma_2\rangle, \langle\sigma_3\rangle]^T$  that specifies the state of polarization of light. Now the locus of the endpoint of this vector is a three-dimensional curve that provides a useful graphical representation of the evolution of the state of polarization of the light as it propagates through a linear optical medium. The Mueller formalism may be regarded as an operation that maps points, representing the input polarization state,  $M_i(\chi_i, \psi_i)$  on the Poincaré sphere  $\Sigma_2^1$  (for pure states) or inside the unit ball  $\Sigma_3^1$  (for mixed states) to new positions  $M_o(\chi_o, \psi_o)$  representing the output polarization state.

Let the surface be  $\Sigma_2^P$ , which defines the states of polarization where the degree of polarization is  $P$ . Of particular importance is to answer the following question: What is the transformation of the surface  $\Sigma_2^{P_i}$ , where  $P_i = |\mathbf{OM}_i|$  is the degree of polarization of input Stokes vectors, by a linear interaction characterized by a general Mueller matrix  $\mathbf{M}$ . In view of our discussion of the Stokes vector space description of partially polarized light in Section 3.2.1,



**Figure 4.1.10.** (a) The surface  $\Sigma_2^{P_i}$ , which defines the states of polarization where the incident degree of polarization is  $P_i$ . The point  $O$  is the origin of the unit ball  $\Sigma^3$ ; (b) the deformation of  $\Sigma_2^{P_i}$  by a general Mueller matrix yields an ellipsoid.

the problem shifts to the evaluation of the length of the vector  $\mathbf{OM}_o$ ,  $|\mathbf{OM}_o|$ , where the point  $M_o$  represents the output polarization state. A proof that  $a \leq |\mathbf{OM}_o| \leq b$ , with  $a$  (resp.  $b$ ) being the lower (resp. upper) bound, is given in Appendix G. Consequently, it turns out that a general Mueller matrix deforms the surface  $\Sigma_2^{P_i}$  into an ellipsoid (see Fig. 4.1.10). A specific example of this deformation by a diagonal Mueller matrix will be discussed later in Section 4.1.5.

We next restrict our consideration to two simple nondepolarizing and nonimaging forming optical devices, namely, the linear polarizer  $\mathbf{P}(\alpha)$  and the

linear compensator  $\mathbf{C}(\delta, \gamma)$ . The evolution of a pure state of polarization of light passing through a nondepolarizing optical device will be represented by a geometric path over the Poincaré sphere  $\Sigma_1^2$ . In accordance with our previous discussion, these polarization mapping properties of optical systems are consistent with the fact that  $SU(2)$  transformations act on  $\Sigma_1^2$  as  $SO(3)$  rotations. From the group theoretical point of view, this parametrization arises from the homomorphic relationship between the groups  $SU(2)$  and  $SO(3)$ .

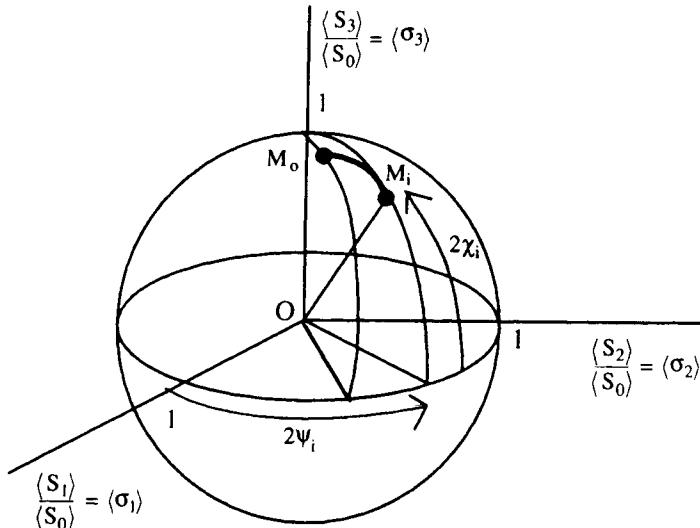
Turn first to examine the behavior of an ideal linear polarizer oriented at angle  $\alpha$ . We consider an incident light in the form of a pure state whose Stokes vector

$$\mathbf{S}_i = \langle S_0 \rangle_i \begin{bmatrix} 1 \\ \mathbf{u}(\chi_i, \psi_i) \end{bmatrix}$$

is incident on a linear polarizer at angle  $\alpha$  characterized by a Mueller matrix  $\mathbf{P}(\alpha)$ , Eq. (4.1.119). The output Stokes vector, shown in Fig. 4.1.11, can be calculated by substituting Eq. (4.1.119) into Eq. (4.1.87).

By using spherical trigonometry, it can be shown that the gain of the linear polarizer is given by

$$g = \cos^2 \left( \frac{\sqrt{M_o M_i}}{2} \right) \quad (4.1.131)$$



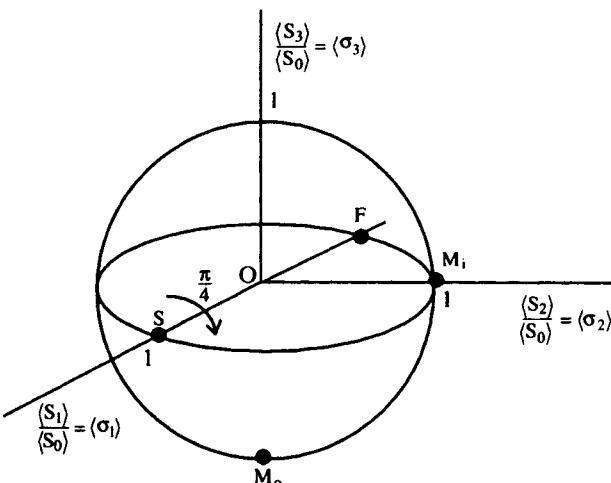
**Figure 4.1.11.** Poincaré sphere representation of the transformation of a polarization state incident on a linear polarizer.

where  $M_o M_i$  denotes the length of the arc joining the two points  $M_o$  and  $M_i$  on  $\Sigma_1^2$  [34]. If  $M_o M_i = \pi$ , that is, if  $M_o$  and  $M_i$  are antipodal states of polarization, no light is transmitted. When the arc varies from 0 to  $\pi$ , the gain decreases from unity (i.e.,  $M_o$  and  $M_i$  are coincident), to zero (i.e.,  $M_o$  and  $M_i$  are orthogonally polarization states). In the case of incident linearly polarized light ( $\chi_i = 0$ ), we recover Malus' law, namely

$$g = \cos^2(\alpha) \quad (4.1.132)$$

The two points  $M_i$  and  $M_o$  that represent respectively the input and output states of polarization lie in the equatorial plane at longitudes differing by  $2\alpha$ .

So far we have been considering the action of a polarizer on a beam of polarized light. We now deal with the behavior of an ideal linear compensator. Input and output elliptical polarizations are points on  $\Sigma_1^2$  that are transformed as a result of phase shifts. The effect of a linear compensator  $C(\delta, \gamma)$  with retardance  $\delta$  and with fast axis at an azimuthal angle  $\theta$  of rotation with respect to the horizontal axis can be described by the transformation  $M_i(\chi_i, \psi_i) \rightarrow M_o(\chi_o, \psi_o)$  whose coordinates can be determined by applying Eqs. (4.1.87) and (4.1.108). Let  $F$  and  $S$  the "fast" and "slow" eigenwaves of the compen-



**Figure 4.1.12.** The geometric path on the Poincaré sphere describes the action of a quarter-wave compensator that transforms an input state of polarization represented by the point  $M_i$  into an output state of polarization represented by the point  $M_o$  by a rotation about the axis  $FS$  through the appropriate angle. The points  $F$  and  $S$  represent the "fast" and "slow" eigenwaves of the compensator, respectively, and are antipodal points.

sator, and let the phase shift between them be  $\delta$ . A compensator's phase shift between fast ( $F$ ) and slow ( $S$ ) axes rotates the input point  $M_i$  about the axis  $FS$  through the angle  $\delta$ . A positive phase shift is a clockwise rotation about this axis is equal to the phase shift. As an illustration of the method, consider a linearly polarized at  $45^\circ$  wave that is incident on a quarter-waveplate  $C(\pi/2)$  with slow axis along  $e_1$  and fast axis along  $e_2$ . The emerging wave is right circularly polarized as indicated by the lower pole of the Poincaré sphere  $M_o$  in Fig. 4.1.12.

If a series of  $N$  compensators is used,  $N$  rotations must be performed in succession, about the appropriate axes and through the appropriate angles.

We leave it to the reader to verify that the output and input azimuth and ellipticity angles are related by the following expressions:

$$\sin(2\chi_o) = \sin(2\chi_i) \cos(\delta) + \cos(2\chi_i) \sin(2(\theta - \psi_i)) \sin(\delta) \quad (4.1.133a)$$

$$\tan(2\psi_o) =$$

$$\frac{\sin(4\theta) \cos(2\psi_i) \sin^2\left(\frac{\delta}{2}\right) + \cos(2\theta) \tan(2\chi_i) \sin(\delta) + \sin(2\psi_i)(\sin^2(2\theta) + \cos^2(2\theta) \cos(\delta))}{\sin(4\theta) \sin(2\psi_i) \sin^2\left(\frac{\delta}{2}\right) - \sin(2\theta) \tan(2\chi_i) \sin(\delta) + \cos(2\psi_i)(\cos^2(2\theta) + \sin^2(2\theta) \cos(\delta))} \quad (4.1.133b)$$

#### 4.1.2.6. Mueller Calculus and Entropy Transformations

The problem of entropy transformation by a linear system was already discussed within the Jones formalism. It remains to be extended here. The calculations are more involved because of the 16 elements that characterize a general Mueller matrix. However, symmetry conditions inherent in the nature of the system or/and certain optical properties may reduce the number of independent parameters; some elements of  $M$  could be equal to others, and some might be zero (a point to which we will return later). This renders the evaluation of the polarization entropy much more tractable. In this case the easier way to compute the entropy is to solve the output Stokes vector using Eq. (4.1.87) and substitution into Eq. (3.4.1).

When the field is Gaussian distributed, we find that the expression of the entropy variation  $\Delta S_{4g}$  is identical to the result obtained with the two-dimensional case,  $\Delta S_2$ ; this is expected since the coherency matrix formalism and covariance matrix of the Stokes parameters are equivalent methods of Gaussian field description [28]. When  $P_o$  take values in the range  $(P_i, 1)$ , corresponding to a nondepolarizing transformation, then  $\Delta S_{4g} \leq 0$  as expected, since this case corresponds to the so-called Mueller–Jones calculus. A special case is again when the matrix  $M$  is unitary. On the other hand, when  $P_o$  take values between 0 and  $P_i$ , corresponding to a depolarizing transformation, then

we can write down an expression for  $\Delta S_{4g}$  that takes positive value. Other aspects of this problem are discussed at greater length elsewhere [35].

#### 4.1.3. RELATIONSHIP BETWEEN JONES AND MUELLER MATRICES

We now are ready to take stock of the similarities and differences between the Jones and Mueller matrix formalisms. Before turning to mathematical details, several caveats are in order. One of the most appealing features of the Mueller matrix is that all its elements are directly measurable. As shown above, the Jones approach is not suitable for extracting information associated with depolarizing effects. A serious limitation of the Mueller calculus is the lack of physical interpretation of all four eigenvalues and eigenvectors [29]. The matrix elements of the former are derived from first principles, that is, Maxwell's equations, whereas the Mueller matrix arises from the assumption that the Stokes vector emerging from the optical system is linearly related to the Stokes vector of the incident wave. The physical parameters of the Jones matrix formalism are amplitudes and phases; those of the Mueller formalism are intensities. The former preserves information to the absolute phase; the latter does not. The former assumes a coherent addition of waves (amplitude); the latter assumes an incoherent addition of waves (intensity). The choice of a particular transformation matrix may be motivated by necessity rather than convenience.

In what follows, we raise the important question: "If we know only the Jones matrix, what can be said about the Mueller matrix and conversely, what are the constraints that must be satisfied for a Mueller matrix to correspond to a Jones matrix?" We will denote by  $M_J$  the Mueller matrix that can be derived from the Jones matrix  $J$  and call it the *Mueller–Jones matrix*. This duality, far from being accidental, is deeply rooted and related to the concept of unitary groups [36].

We start by considering an optical system characterized by a Jones matrix  $J$ . Let us define the  $4 \times 1$  complex vector  $F$  as

$$F = \langle E \otimes E^* \rangle = \begin{bmatrix} \Phi_{11} \\ \Phi_{12} \\ \Phi_{21} \\ \Phi_{22} \end{bmatrix} \quad (4.1.134)$$

Note that the elements of  $F$ , which are identical to those of  $\Phi_2$ , have the same physical significance. The Stokes vector may be expressed from  $F$  by making use of the unitary transformation

$$S = AF \quad (4.1.135)$$

where the constant matrix  $\mathbf{A}$  ( $2\mathbf{A}^{-1} = \mathbf{A}^+ = (\mathbf{A}^*)^T$ ) is given by

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \quad (4.1.136)$$

On substituting Eq. (4.1.1) into Eq. (4.1.134), and if we further use a standard property of Kronecker product of matrices, we obtain the following expression for the output complex vector  $\mathbf{F}_o$ :

$$\mathbf{F}_o = \langle \mathbf{E}_o \otimes \mathbf{E}_o^* \rangle = \langle \mathbf{J}\mathbf{E}_i \otimes \mathbf{J}^*\mathbf{E}_i^* \rangle = \langle (\mathbf{J} \otimes \mathbf{J}^*)(\mathbf{E}_i \otimes \mathbf{E}_i^*) \rangle = (\mathbf{J} \otimes \mathbf{J}^*)\mathbf{F}_i \quad (4.1.137)$$

This, in turn, can be written as

$$\mathbf{S}_o = \mathbf{A}\mathbf{F}_o = (\mathbf{J} \otimes \mathbf{J}^*)\mathbf{A}^{-1}\mathbf{A}\mathbf{F}_i = \mathbf{M}_j\mathbf{S}_i = (\mathbf{A}(\mathbf{J} \otimes \mathbf{J}^*)\mathbf{A}^{-1})\mathbf{S}_i \quad (4.1.138)$$

By virtue of Eq. (4.1.87), we arrive at the relation

$$\mathbf{M}_j = \mathbf{A}(\mathbf{J} \otimes \mathbf{J}^*)\mathbf{A}^{-1} \quad (4.1.139)$$

This formula is sufficiently interesting to warrant some comment. Since Eq. (4.1.139) is invariant under the transformation  $\mathbf{J} \rightarrow \mathbf{J} \exp(i\varphi)$ , one can calculate  $\mathbf{J}$  from  $\mathbf{M}_j$  only within an arbitrary phase factor  $\varphi$ . Moreover, it can be verified that if  $\mathbf{M}_j$  is a Mueller–Jones matrix, then so is  $\mathbf{M}_j^T$ . As discussed in texts on group theory, these matrices constitute a representation of the group of proper homogeneous Lorentz transformations in space–time [10].

A number of properties follow from the definition (4.1.139) if we make use of basic properties of Kronecker products. Let  $\mathbf{M}_{jk}$  be the Mueller–Jones matrices with associated Jones matrices  $\mathbf{J}_k$ , respectively; then the reader can prove by induction the following theorems [38]. If

$$\mathbf{M}_j = \prod_{k=N}^1 \mathbf{M}_{jk} \quad \text{then} \quad \mathbf{J} = \prod_{k=N}^1 \mathbf{J}_k \quad (4.1.140a)$$

If

$$|\det(\mathbf{J})| \neq 0 \quad \text{and} \quad \mathbf{M}_j \rightarrow \mathbf{M}_j^{-1} \quad \text{then} \quad \mathbf{J} \rightarrow \mathbf{J}^{-1} \quad (4.1.140b)$$

If

$$\mathbf{M}_j \rightarrow \alpha \mathbf{M}_j \quad \text{then} \quad \mathbf{J} \rightarrow \sqrt{\alpha} \mathbf{J} \quad (4.1.140c)$$

$$\det(\mathbf{M}_j) \geq 0 \quad (4.1.140d)$$

where  $\alpha$  is a positive scalar.

Assuming the most general form of  $\mathbf{J}$  given by Eq. (4.1.2), the explicit form of the  $\mathbf{M}_J$  reads

$$\mathbf{M}_J = \begin{bmatrix} \frac{1}{2}(|J_{11}|^2 + |J_{12}|^2 + |J_{21}|^2 + |J_{22}|^2) & \frac{1}{2}(|J_{11}|^2 - |J_{12}|^2 + |J_{21}|^2 - |J_{22}|^2) & \operatorname{Re}(J_{11}^* J_{12} + J_{21}^* J_{22}) & -\operatorname{Im}(J_{11}^* J_{12} + J_{21}^* J_{22}) \\ \frac{1}{2}(|J_{11}|^2 + |J_{12}|^2 - |J_{21}|^2 - |J_{22}|^2) & \frac{1}{2}(|J_{11}|^2 - |J_{12}|^2 - |J_{21}|^2 + |J_{22}|^2) & \operatorname{Re}(J_{11}^* J_{12} - J_{21}^* J_{22}) & -\operatorname{Im}(J_{11}^* J_{12} - J_{21}^* J_{22}) \\ \operatorname{Re}(J_{11}^* J_{21} + J_{12}^* J_{22}) & \operatorname{Re}(J_{11}^* J_{21} - J_{12}^* J_{22}) & \operatorname{Re}(J_{11}^* J_{22} + J_{12}^* J_{12}) & \operatorname{Im}(J_{11}^* J_{22} - J_{12}^* J_{12}) \\ \operatorname{Im}(J_{11}^* J_{21} + J_{12}^* J_{22}) & \operatorname{Re}(J_{11}^* J_{21} - J_{12}^* J_{22}) & \operatorname{Im}(J_{11}^* J_{22} + J_{12}^* J_{12}) & \operatorname{Re}(J_{11}^* J_{22} - J_{12}^* J_{12}) \end{bmatrix} \quad (4.1.141)$$

We can also derive a second expression of  $\mathbf{M}_J$ . Consider Eq. (4.1.7), which can be rewritten with the help of Eq. (3.1.130) as

$$\sum_{l=0}^3 \langle S_l \rangle_o \sigma_l = \sum_{l=0}^3 \langle S_l \rangle_i \mathbf{J} \boldsymbol{\sigma}_l \mathbf{J}^+ \quad (4.1.142)$$

Multiply both sides by  $\boldsymbol{\sigma}_k$  and take the trace. Then, making use of the trace orthogonality of the Pauli matrices, yields

$$2 \langle S_l \rangle_o = \sum_{l=0}^3 \langle S_l \rangle_i \operatorname{tr}(\mathbf{J} \boldsymbol{\sigma}_i \mathbf{J}^+ \boldsymbol{\sigma}_k) = \sum_{l=0}^3 \langle S_l \rangle_i \mu_{kl} \quad (4.1.143)$$

Thus we find the desired result:

$$(M_J)_{kl} = \mu_{kl} = \frac{1}{2} \operatorname{tr}(\boldsymbol{\sigma}_k \mathbf{J} \boldsymbol{\sigma}_l \mathbf{J}^+) \quad (4.1.144)$$

Since a Jones matrix is determined only up to overall phase, Mueller–Jones matrices can contain up to seven independent matrix elements (if an inessential absolute phase is not included), so there are nine relations between the 16 matrix elements of  $\mathbf{M}_J$  [39]. It should be noted that for every Jones matrix  $\mathbf{J}$  there is a corresponding Mueller matrix  $\mathbf{M}_J$ , but the converse is not true. This is to be expected since the Jones matrix contains fewer degrees of freedom than does a Mueller matrix. Our goal in this analysis is to derive conditions that a Mueller matrix be a Mueller–Jones matrix. By multiplying Eq. (4.1.139) by its transpose ( $(\mathbf{M}_J^T \equiv \mathbf{M}_J^+$  since  $\mathbf{M}_J$  is real-valued), we arrive at the relation

$$\begin{aligned} \mathbf{M}_J^T \mathbf{M}_J &= (\mathbf{A}(\mathbf{J} \otimes \mathbf{J}^*) \mathbf{A}^{-1})^+ \mathbf{A}(\mathbf{J} \otimes \mathbf{J}^*) \mathbf{A}^{-1} \\ &= \frac{1}{2} \mathbf{A}(\mathbf{J} \otimes \mathbf{J}^*)^+ \mathbf{A}^+ \mathbf{A}(\mathbf{J} \otimes \mathbf{J}^*) \mathbf{A}^{-1} \end{aligned} \quad (4.1.145)$$

If we now use the general matrix rules

$$\operatorname{tr}(\mathbf{C} \otimes \mathbf{D}) = \operatorname{tr}(\mathbf{C}) \operatorname{tr}(\mathbf{D}) \quad (4.1.146a)$$

$$\mathbf{CD} \otimes \mathbf{EF} = (\mathbf{C} \otimes \mathbf{E})(\mathbf{D} \otimes \mathbf{F}) \quad (4.1.146b)$$

$$(\mathbf{C} \otimes \mathbf{D})^+ = \mathbf{C}^+ \otimes \mathbf{D}^+ \quad (4.1.146c)$$

where  $\mathbf{C}$ ,  $\mathbf{D}$ ,  $\mathbf{E}$ ,  $\mathbf{F}$  are four arbitrary square matrices, Eq. (4.1.145) reduces to

$$\mathbf{M}_J^T \mathbf{M}_J = \mathbf{A}(\mathbf{J} \otimes \mathbf{J}^*)^+ (\mathbf{J} \otimes \mathbf{J}^*) \mathbf{A}^{-1} = \mathbf{A}(\mathbf{J}^+ \mathbf{J}) \otimes (\mathbf{J}^+ \mathbf{J})^* \mathbf{A}^{-1} \quad (4.1.147)$$

On taking the trace of both sides and making use of the invariance of the trace under cyclic permutations, we obtain

$$\text{tr}(\mathbf{M}_J^T \mathbf{M}_J) = (\text{tr}(\mathbf{J}^+ \mathbf{J}))^2 \quad (4.1.148)$$

It follows at once from Eq. (4.1.144) that  $\mu_{00} = \frac{1}{2}(\text{tr}(\mathbf{J}^+ \mathbf{J}))$ ; hence Eq. (4.1.148) reduces to

$$\text{tr}(\mathbf{M}_J^T \mathbf{M}_J) = \sum_{i,j=0}^3 \mu_{ij}^2 = 4\mu_{00}^2 \quad (4.1.149)$$

which is termed the *trace condition*. We should note that the result quoted in Eq. (4.1.149) was first proposed by Gil and Bernabeu [40]. It is a necessary condition that a Mueller matrix be equivalent to a Jones matrix. However, as was pointed out in Refs 30 and 39, with explicit counterexamples, Eq. (4.1.149) is not sufficient. Note also that if  $\mathbf{J}$  is a unitary matrix,  $\mathbf{M}_J$  is also unitary and conversely. The proof follows directly from Eq. (4.1.139) and from the fact that the general solution of the equation  $\mathbf{O} \otimes \mathbf{O}^* = \mathbf{O}_0^{(4)}$ , where  $\mathbf{O}_0^{(4)}$  is the  $4 \times 4$  unit matrix, is  $\mathbf{O} = \exp(i\Delta)\sigma_0$ . The specific role of  $\mu_{00}$  in Eq. (4.1.149) can be recognized as being the gain for unpolarized light [Eq. (4.1.95)]. Anderson and Barakat have generalized Eq. (4.1.149) and found 16 necessary conditions that are required for a given Mueller matrix to become a Jones matrix

$$\text{tr}((\boldsymbol{\sigma}_j \otimes \boldsymbol{\sigma}_0) \mathbf{H} (\boldsymbol{\sigma}_k \otimes \boldsymbol{\sigma}_0) \mathbf{H}^*) = 4\mu_{00}\mu_{jk} \quad (4.1.150)$$

where  $\mathbf{H} \equiv \mathbf{A}^{-1} \mathbf{M}_J \mathbf{A}$  and  $\mathbf{A}$  is given by Eq. (4.1.136) [42]. As the reader can see, Eq. (4.1.150) includes Eq. (4.1.149) as the special case  $j = k = 0$ .

As might be expected, the gain of a Mueller–Jones matrix must reduce to the gain of the corresponding Jones matrix. It follows directly from Eq. (4.1.94) that we can write

$$g = \mu_{00} + \sum_{j=1}^3 \mu_{0j} \langle \boldsymbol{\sigma}_j \rangle \quad (4.1.151)$$

Typical examples of Mueller–Jones matrices include polarizers, compensators,

and rotators. The depolarizing filter

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

is an example of Mueller matrix that is not derivable from a Jones matrix and hence is not a Mueller–Jones matrix.

At this point it is useful to examine some consequences of Eq. (4.1.91). Taking first the determinant of both sides leads to

$$\det(\Phi_{4o}) = \det(\Phi_{4i}) |\det(\mathbf{M}_J)|^2 \quad (4.1.152)$$

This expression generalizes Eq. (4.1.9) obtained via the  $2 \times 2$  coherency matrix formalism. Using Eqs. (4.1.9) and (3.1.142), we find that

$$\langle \mathbf{S} \rangle_o^T \mathbf{G} \langle \mathbf{S} \rangle_o = |\det(\mathbf{J})|^2 \langle \mathbf{S} \rangle_i^T \mathbf{G} \langle \mathbf{S} \rangle_i \quad (4.1.153)$$

where  $\mathbf{G}$  is the Lorentz metric matrix, Eq. (3.1.143). Substitution of Eq. (4.1.139) into Eq. (4.1.153) yields

$$(1 - P_o^2) = \frac{\det(\mathbf{M}_J)}{g^2} (1 - P_i^2) \quad (4.1.154)$$

which parallels Eq. (4.1.11). On substituting  $\langle \mathbf{S} \rangle_o = \mathbf{M}_J \langle \mathbf{S} \rangle_i$  into Eq. (4.1.153), we arrive at the following expression:

$$\mathbf{M}_J^T \mathbf{G} \mathbf{M}_J = |\det(\mathbf{J})|^2 \mathbf{G} = (\det(\mathbf{M}_J))^{1/2} \mathbf{G} \quad (4.1.155)$$

which is satisfied by any deterministic Mueller–Jones matrix. It is worth noting that Eq. (4.1.155) is at the basis of the derivation of the nine bilinear constraints between the elements of the Mueller–Jones matrix [43]. The following nine formulas can be derived:

$$\mu_{00}\mu_{01} - \mu_{10}\mu_{11} - \mu_{20}\mu_{21} - \mu_{30}\mu_{31} = 0 \quad (4.1.156a)$$

$$\mu_{00}\mu_{02} - \mu_{10}\mu_{12} - \mu_{20}\mu_{22} - \mu_{30}\mu_{32} = 0 \quad (4.1.156b)$$

$$\mu_{00}\mu_{03} - \mu_{10}\mu_{13} - \mu_{20}\mu_{23} - \mu_{30}\mu_{33} = 0 \quad (4.1.156c)$$

$$\mu_{01}\mu_{02} - \mu_{11}\mu_{12} - \mu_{21}\mu_{22} - \mu_{31}\mu_{32} = 0 \quad (4.1.156d)$$

$$\mu_{01}\mu_{03} - \mu_{11}\mu_{13} - \mu_{21}\mu_{23} - \mu_{31}\mu_{33} = 0 \quad (4.1.156e)$$

$$\mu_{02}\mu_{03} - \mu_{12}\mu_{13} - \mu_{22}\mu_{23} - \mu_{32}\mu_{33} = 0 \quad (4.1.156f)$$

$$\mu_{01}^2 - \mu_{11}^2 - \mu_{21}^2 - \mu_{31}^2 + \mu_{00}^2 - \mu_{10}^2 - \mu_{20}^2 - \mu_{30}^2 = 0 \quad (4.1.156g)$$

$$\mu_{02}^2 - \mu_{12}^2 - \mu_{22}^2 - \mu_{32}^2 + \mu_{00}^2 - \mu_{10}^2 - \mu_{20}^2 - \mu_{30}^2 = 0 \quad (4.1.156h)$$

$$\mu_{03}^2 - \mu_{13}^2 - \mu_{23}^2 - \mu_{33}^2 + \mu_{00}^2 - \mu_{10}^2 - \mu_{20}^2 - \mu_{30}^2 = 0 \quad (4.1.156i)$$

The relations (4.1.156a–i) can be inverted. We have

$$|J_{11}|^2 = \frac{1}{2}(\mu_{00} + \mu_{01} + \mu_{10} + \mu_{11}) \quad (4.1.157a)$$

$$|J_{12}|^2 = \frac{1}{2}(\mu_{00} - \mu_{01} + \mu_{10} - \mu_{11}) \quad (4.1.157b)$$

$$|J_{21}|^2 = \frac{1}{2}(\mu_{00} + \mu_{01} - \mu_{10} - \mu_{11}) \quad (4.1.157c)$$

$$|J_{22}|^2 = \frac{1}{2}(\mu_{00} - \mu_{01} - \mu_{10} + \mu_{11}) \quad (4.1.157d)$$

$$\cos(\Omega_{11} - \Omega_{12}) = \frac{(\mu_{02} + \mu_{12})}{2|J_{11}||J_{12}|} \quad (4.1.157e)$$

$$\cos(\Omega_{21} - \Omega_{22}) = \frac{(\mu_{02} - \mu_{12})}{2|J_{21}||J_{22}|} \quad (4.1.157f)$$

$$\cos(\Omega_{11} - \Omega_{21}) = \frac{(\mu_{20} + \mu_{21})}{2|J_{11}||J_{21}|} \quad (4.1.157g)$$

$$\cos(\Omega_{12} - \Omega_{22}) = \frac{(\mu_{20} - \mu_{21})}{2|J_{12}||J_{22}|} \quad (4.1.157h)$$

$$|J_{11}||J_{22}|\cos(\Omega_{11} - \Omega_{22}) + |J_{12}||J_{21}|\cos(\Omega_{12} - \Omega_{21}) = \mu_{22} \quad (4.1.157i)$$

where  $|J_{ij}|$  and  $\Omega_{ij}$  are the absolute value and angle of  $J_{ij}$  [i.e.,  $J_{ij} = |J_{ij}| \exp(i\Omega_{ij})$ ]. Only differences between the angles of the Jones matrix elements can be determined by Eqs. (4.1.157a–i), in accordance with the fact that the matrix  $\mathbf{J}$  is determined only up to overall phase. These equations allow an experimenter to recover the Jones matrix from the Mueller description. We note finally that Anderson and Barakat have shown that any Mueller matrix can be expressed as a linear combination of at most four trace orthonormal Mueller–Jones matrices [42].

#### 4.1.4. POLARIZATION TRANSFER IN NONDEPOLARIZING OPTICAL LINEAR MEDIA

In the discussion of the coherency (density) matrix formalism, we found that a quantity of considerable interest in the study of partially polarized light is the degree of polarization of a Stokes vector. The effect of a general Mueller matrix on the degree of polarization of an incident Stokes vector can be complicated; it may increase, maintain, or decrease the degree of polarization. The present section is devoted to a derivation of the conditions that a Mueller optical

medium cannot decrease the degree of polarization for any incident Stokes vector [41]. Mathematically it translates into finding the output degree of polarization and comparing it to the degree of polarization of the incident wave.

We have already seen that the output normalized Stokes parameters can be expressed as real linear combinations of the input normalized Stokes parameters via Eq. (4.1.87). We have

$$\langle \sigma_j \rangle_o = \frac{1}{g} \sum_{l=0}^3 m_{jl} \langle \sigma_l \rangle_0 \quad (4.1.158)$$

We next substitute this expression in the definition of the output degree of polarization, Eq. (3.1.150a). The result is

$$1 + P_o^2 = \frac{1}{g^2} \left( \sum_{l=0}^3 \alpha_l \langle \sigma_l \rangle_0^2 + 2 \sum_{l=0}^3 \sum_{p>l} \beta_{lp} \langle \sigma_l \rangle_0 \langle \sigma_p \rangle_0 \right) \quad (4.1.159)$$

where  $\alpha_{ll} = (\mathbf{M}^T \mathbf{M})_{ll} \geq 0$  and  $\beta_{lp} = (\mathbf{M}^T \mathbf{M})_{lp} = \beta_{pl}$ . Next, from Eq. (4.1.159), we can derive the following expression for  $P_o^2 - P_i^2$ :

$$P_o^2 - P_i^2 = \sum_{l,p=0}^3 \left( \frac{1}{g^2} \beta_{lp} - \delta_{lp} \right) \langle \sigma_l \rangle_0 \langle \sigma_p \rangle_0 = Q = \sum_{l,p=0}^3 a_{lp} \langle \sigma_l \rangle_0 \langle \sigma_p \rangle_0 \quad (4.1.160)$$

It is readily verifiable that  $Q$  is a real positive semidefinite quadratic form, where  $a_{lp}$  are the elements of the matrix  $\mathbf{L} = \mathbf{M}^T \mathbf{M} / g^2 - \mathbf{O}_0^{(4)}$ . According to the Rayleigh principle, an optimization of the quadratic form  $Q$  leads to an eigenvalue problem [21]. Following this analysis, one then obtains

$$\inf \sum_{l=0}^3 |\langle \sigma_l \rangle|^2 \leq Q \leq \sup \sum_{l=0}^3 |\langle \sigma_l \rangle|^2 \quad (4.1.161)$$

where  $\inf$  (resp.  $\sup$ ) is the smallest (resp. largest) eigenvalue of matrix  $\mathbf{L}$ . Since  $1 \leq \sum_{l=0}^3 |\langle \sigma_l \rangle|^2 \leq 2$ , we have

$$\inf \leq P_o^2 - P_i^2 \leq 2 \sup \quad (4.1.162)$$

Denoting by  $\beta(\mathbf{M}^T \mathbf{M})$  the eigenvalues of the matrix  $\mathbf{M}^T \mathbf{M}$ , we have  $\beta(\mathbf{L}) = \beta(\mathbf{M}^T \mathbf{M}) / g^2 - 1$ , and we find that the Eq. (4.1.162) is satisfied if

$$|\inf| \leq 1 \quad (4.1.163a)$$

and

$$0 \leq \sup \leq \frac{\text{tr}(\mathbf{M}^T \mathbf{M})}{(g_{\min})^2} - 4 \quad (4.1.163b)$$

where  $g_{\min}$  is the minimum value of the transmittance, Eq. (4.1.18b) [41].

As an example to appreciate the significance of Eqs. (4.1.163a, b), we consider the special case of unitary Mueller matrices,  $\mathbf{M}^T = \mathbf{M}^{-1}$ . In that case, we easily find that  $\inf = \sup = 0$ , which is consistent with the fact that optical systems characterized by unitary Mueller matrices do not affect the degree of polarization of an incident light. Further aspects of Eqs. (4.1.163a, b) have been discussed by several authors. Brosseau has utilized Eqs. (4.1.163a, b) to characterize experimentally determined Mueller matrices [41]. Gerligand and colleagues have used Eqs. (4.1.163a, b) to test Mueller matrix measurements of ferrofluids in magnetic fields [44].

#### 4.1.5. MUELLER MATRIX ANALYSIS OF LIGHT DEPOLARIZATION BY A LINEAR OPTICAL MEDIUM

One focal point of physical optics in recent times is the study of depolarization of light induced by the presence of disorder, such as a suspension of spherical diffusers in liquids and rough and heterogeneous surfaces. At the present time, there is a large collection of Mueller matrix data available in the literature. The refinements of the analytic and numerical tools and the confrontation with experimental data have set the stage for the quest of understanding the physical mechanisms that cause the depolarization. At least two mechanisms have been proposed to explain this phenomenon: (1) decorrelation of the phases and the amplitudes of the electric field vector and (2) selective absorption of polarization states. An experimentally determined or computed Mueller matrix contains all information concerning the optical properties of the system. With no real consensus on how to define the term *depolarization*, it is not surprising to find a great variety of definitions in the literature, and this can lead to considerable confusion to those who are new to the field and consternation to the old. Generally speaking, it seems that the term *depolarization* has come to mean the decrease of the degree of polarization of an incident light beam. For a nondepolarizing optical system, the degree of polarization of an incident light  $P_i$  is always less than, or equal to, the degree of polarization  $P_o$  of the light emerging at the output of this optical system whatever the particular state of polarization of the incident light. This definition is the most common usage of the term *depolarization* and is the definition retained here [24]. At this stage of the discussion it is also worth pointing out that the problem of depolarization is not the same physical situation of the Mueller–Jones equivalence issue as discussed in Section 4.1.4.

The purpose of this section is to analyze depolarization of light by a linear optical medium using the Mueller matrix formalism. This is an important question in practice because experimentalists need methods to discriminate between depolarizing and nondepolarizing parts of a Mueller matrix, and this for a wide variety of applications, including calibration of polarimetric instruments, and polarization signatures of targets embedded in clutter. We develop a general strategy, which we suggest has applications in the extraction of physical information from a range of polarization matrices.

If we restrict consideration to incident Stokes vectors in the form of pure states ( $P_i = 1$ ), which is usually the case encountered in practice for measuring a Mueller matrix, we can classify the Mueller polarization matrices as follows. Three cases can occur.

1. If  $P_o = 1$  for all incident states of polarization  $S_i$ , Eq. (4.1.139) is satisfied,  $\mathbf{M}$  is derivable from a Jones matrix, and thence is a Mueller–Jones matrix  $\mathbf{M}_J$ . A key point to be stressed here is that characterization of depolarization, by this particular optical system, of light in the form of mixed states can be done using the polar decomposition of  $\mathbf{J}$  [Eqs. (4.1.66a, b)]. Unitary matrices are of particular interest because they let invariant the gain and the degree of polarization (e.g., pure rotation). Hence the depolarization behavior is entirely contained in the Hermitian matrix  $\mathbf{H}$ .
2. If  $P_o < 1$  for at least one incident state of polarization  $S_i$ , we can decompose  $\mathbf{M}$  as a sum of a part satisfying Eq. (4.1.139) and a part whose output Stokes vector is proportional to

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\mathbf{M} = (1 - \beta)\mathbf{M}_J + \beta\mathbf{D} \quad (4.1.164)$$

where

$$\beta = m_{00} \left[ 1 - \frac{1}{3} \left( 3 \left( \frac{\text{tr}(\mathbf{M}^T \mathbf{M})}{m_{00}^2} - 1 \right) \right)^{1/2} \right] \quad \text{and} \quad \mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

is the Mueller matrix of an ideal depolarizer. As in the first case, we notice that the polar decomposition of  $\mathbf{J}$  permits to treat the depolarization of partially polarized light [45].

3. The third and final case is when  $P_o < 1$  for all incident states of polarization  $\mathbf{S}_i$ . This last situation corresponds to a pure depolarizer.

Several conclusions are already emerging. An important consequence is that we can predict whether depolarization of pure states of polarization occurs by considering Eq. (4.1.164). At the very least, this amounts to an efficient way of studying the physical mechanisms of light depolarization by a deterministic optical medium. A second key feature of this analysis is the derivation of a formula for a Mueller matrix as a linear combination of a depolarizing and a nondepolarizing parts.

Since our results to this point have all been stated in fairly general terms, it may be of help to discuss several deterministic Mueller matrices reported in the literature.

**Example 4.1.1. Mueller Matrix that Is a Mueller–Jones Matrix.** The following matrix derives from the research of Ramsey concerning a measurement of reflection from a smooth, flat, homogeneous steel surface [46]:

$$\mathbf{M}_1 = \begin{bmatrix} 0.5491 & -0.1982 & -0.0096 & 0.0004 \\ -0.1991 & 0.5492 & 0.0389 & -0.0102 \\ -0.0064 & 0.0372 & -0.4285 & 0.2804 \\ 0.0004 & 0.0111 & -0.2814 & -0.4281 \end{bmatrix} \quad (4.1.165)$$

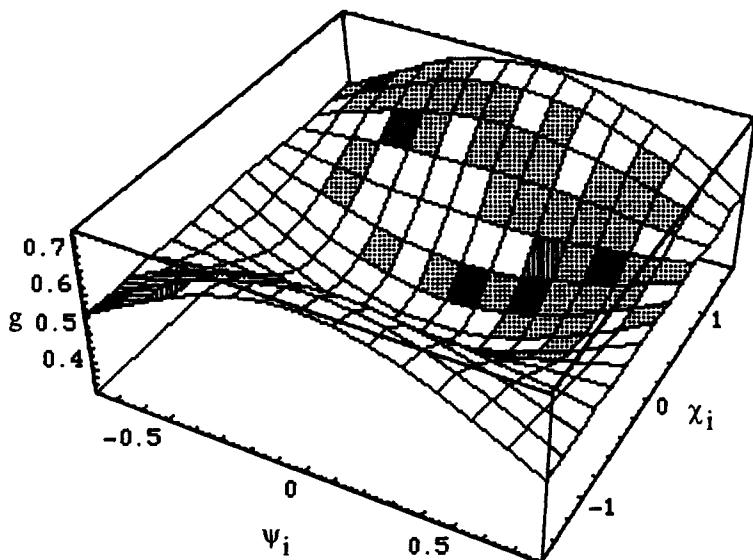
$\mathbf{M}_1$  is a physically realizable Mueller matrix as illustrated in Fig. 4.1.13 for the gain and Fig. 4.1.14 for the degree of polarization. A measure of uncertainty of the elements of this matrix, reported in Ref. 44, gives  $\pm 0.005$ . It is worth noting that we have  $P_o \cong 1$  for all incident states of polarization. This is consistent with the fact that  $\mathbf{M}_1$  is a measurement of reflection from a polished metal and is expected to correspond to an ideal reflector.

On one hand,  $\mathbf{M}_1$  passes the necessary condition Eq. (4.1.149) since we find that  $\text{tr}(\mathbf{M}_1^T \mathbf{M}_1)/4(\mathbf{M}_1)_{00}^2 = 1.0033$ . On the other hand, Eq. (4.1.155) is used to test  $\mathbf{M}_1$ . The result is shown in Table 4.1.2, where the top element is  $\mathbf{M}_1^T \mathbf{G} \mathbf{M}_1$  and the bottom element is  $(\det(\mathbf{M}_1))^{1/2} \mathbf{G}$ .

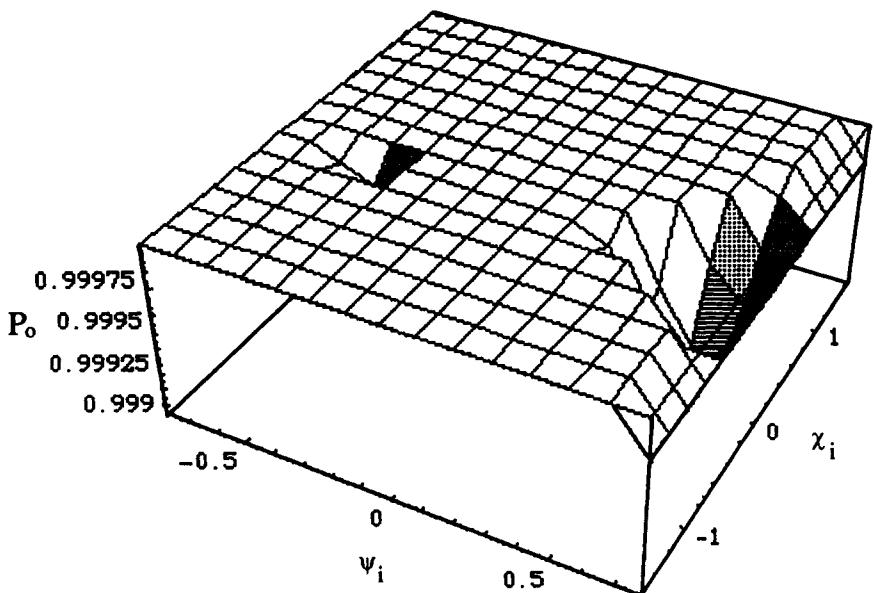
Therefore it is reasonable to conclude that  $\mathbf{M}_1$  is a Mueller–Jones matrix. This fact is confirmed by the recent analysis of Anderson and Barakat [42]. Within the level of uncertainty in the specification of  $\mathbf{M}_1$ , the corresponding Jones matrix may be approximated by

$$\mathbf{J}_1 \cong \begin{bmatrix} 0.5924 & 0 \\ 0 & -0.7231 + i0.4742 \end{bmatrix} \quad (4.1.166)$$

These conclusions are consistent with the fact that  $\mathbf{J}_1$  satisfies the condition for the physical realizability of passive deterministic Jones matrices, given by



**Figure 4.1.13.** Three-dimensional plot of the gain of the optical medium characterized by Mueller matrix  $\mathbf{M}_1$  as a function of the incident state of polarization parametrized by the ellipsometric angles  $\chi_i$  and  $\psi_i$ .



**Figure 4.1.14.** Output degree of polarization of the optical medium characterized by Mueller matrix  $\mathbf{M}_1$  as a function of the incident state of polarization parametrized by the ellipsometric angles  $\chi_i$  and  $\psi_i$ .

**TABLE 4.1.2.** Test (4.1.155) for  $M_t^a$ 

$j/k$	0	1	2	3
0	0.2618	0.0007	-0.0002	0.0002
	0.2630	0	0	0
1	0.0008	-0.2638	-0.0004	-0.0002
	0	-0.2630	0	0
2	-0.0002	-0.0004	-0.2642	0.0001
	0	0	-0.2630	0
3	0.0002	-0.0002	0.0001	-0.2620
	0	0	0	-0.2630

<sup>a</sup>The top element is  $\mathbf{M}_t^T \mathbf{G} \mathbf{M}_t$ , and the bottom element is  $(\det(\mathbf{M}_t))^{1/2} \mathbf{G}$ .

Eq. (4.1.61). These features are consistent with Ramsey's assumption that this measurement corresponds to an ideal reflector described by a Mueller–Jones matrix of the form

$$\mathbf{M} = \begin{bmatrix} a & b & 0 & 0 \\ b & a & 0 & 0 \\ 0 & 0 & c & d \\ 0 & 0 & -d & c \end{bmatrix}$$

Observe that  $\mathbf{M}$  depends on four independent parameters, leaving three degrees of freedom (up to overall phase) for the corresponding Jones matrix, Eq. (4.1.66). Testing Eqs. (4.1.163a, b), we have

$$|\inf| = -0.78 < 1 \quad \text{and} \quad 0 < \sup = 3.58 < \frac{\text{tr}(\mathbf{M}_J^T \mathbf{M}_J)}{(g_{\min})^2} - 4 = 5.93$$

We therefore judge, within the experimental uncertainties, that  $\mathbf{M}_t$  is a measurement of a Mueller–Jones matrix that cannot decrease the degree of polarization of an incident partially polarized wave.

**Example 4.1.2. Mueller Matrix that Can Be Decomposed According to Eq. (4.1.164).** Now the Mueller matrix that we consider is taken from a measurement of reflection from a dielectric underwater target [31]. As was proved in Section 4.1.2.3,  $\mathbf{M}_2$  is a physically realizable Mueller matrix.

Testing Eqs. (4.1.149) and (4.1.155) indicates that  $\mathbf{M}_2$  is not a Mueller–Jones matrix. From Eq. (4.1.149), we have  $\text{tr}(\mathbf{M}_2^T \mathbf{M}_2)/4(\mathbf{M}_2)_{00}^2 = 0.4257 < 1$ . This is consistent with the data of Table 4.1.3 showing a comparison of the right and left members of Eq. (4.1.155), where significant deviations can be observed. The corresponding Jones matrix assuming  $\mathbf{M}_2$  to be of the form of

TABLE 4.1.3. Test (4.1.155) for  $\mathbf{M}_2^a$ 

$j/k$	0	1	2	3
0	0.9992	-0.0229	0.0091	-0.0004
	0.3350	0	0	0
1	-0.0229	-0.2443	-0.0069	-0.0067
	0	-0.3350	0	0
2	0.0091	-0.0069	-0.2555	-0.0032
	0	0	-0.3350	0
3	-0.0004	-0.0067	-0.0032	-0.2018
	0	0	0	-0.3350

<sup>a</sup>The top element is  $\mathbf{M}_2^T \mathbf{G} \mathbf{M}_2$ , and the bottom element is  $(\det(\mathbf{M}_2))^{1/2} \mathbf{G}$ .

Eq. (4.1.164), with  $\beta = 0.516$ , can be obtained in a manner identical to that I used for  $\mathbf{M}_1$ .

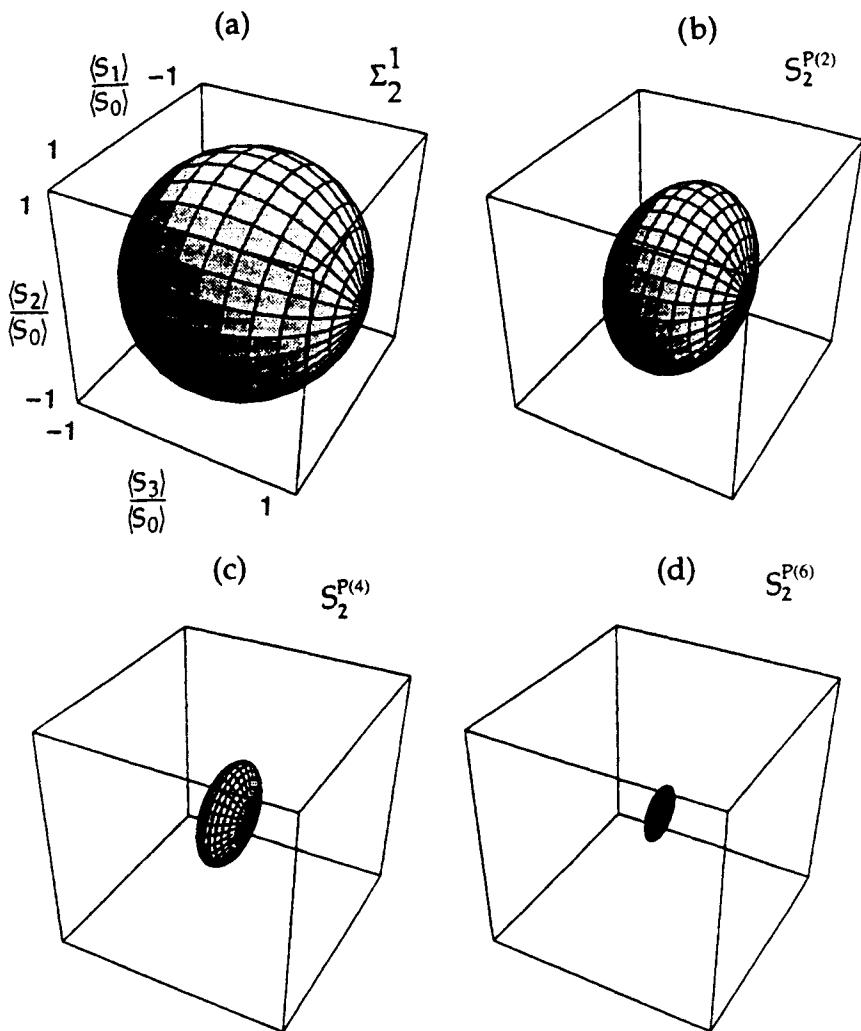
**Example 4.1.3. Mueller Matrix that Cannot Be Decomposed According to Eq. (4.1.164).** Let us finally examine an example of the set of Mueller matrices that are not derivable from Jones matrices. It was pointed out by Bicout and Brosseau [47], based on a maximum entropy principle (MEP), that the Mueller matrix arising in multiple scattering of light by spherical pointlike diffusers (Rayleigh regime) is given by

$$\mathbf{M}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & m_{11} & 0 & 0 \\ 0 & 0 & m_{11} & 0 \\ 0 & 0 & 0 & m_{33} \end{bmatrix} \quad (4.1.167)$$

where  $m_{11}(n) = 3(\frac{7}{10})^n / [2 + (\frac{7}{10})^n]$  and  $m_{33}(n) = 3(\frac{1}{2})^n / [2 + (\frac{7}{10})^n]$  where  $n + 1$  is the number of scattering events. A discussion of the physical arguments leading to Eq. (4.1.167) is postponed until Section 4.4. Again we start by testing Eqs. (4.1.149) and (4.1.167). When  $n \geq 1$ , neither Eqs. (4.1.149) nor (4.1.155) is satisfied; we conclude that  $\mathbf{M}_3$  is not a Mueller-Jones matrix. Observe that  $\mathbf{M}_3$  depolarizes any arbitrary incident Stokes vector (i.e.,  $\lim_{n \rightarrow \infty} \mathbf{M}_3 = \mathbf{D}$ ); thence we can assert that  $\mathbf{M}_3$  cannot be decomposed according to Eq. (4.1.164).

Finally we consider the surface  $S_2^{P(n)}$  parametrized by the polar coordinates of the output Stokes vector

$$\mathbf{S}_o \equiv \begin{bmatrix} m_{11}(n) \cos(2\chi_i) \cos(2\psi_i) \\ m_{11}(n) \cos(2\chi_i) \sin(2\psi_i) \\ m_{33}(n) \sin(2\psi_i) \end{bmatrix}$$



**Figure 4.1.15.** Parametric plot of the surface  $S_2^{P(n)}$  for different values of the number of scattering events  $n$ . Rayleigh scattering by pointlike dielectric spheres: (a)  $n = 0$ , (b)  $n = 2$ , (c)  $n = 4$ , (d)  $n = 6$ .

with  $\chi_i$  running from  $-(\pi/2)$  to  $\pi/2$  and  $\psi_i$  from  $-(\pi/4)$  to  $\pi/4$ . For  $n > 1$ , the process of depolarization cannot be assimilated to an isotropic contraction of the Poincaré sphere  $\Sigma_1^2$  but induces a symmetry breaking (i.e., the symmetry of  $SO(3)$  is broken). For scattering by a single particle ( $n = 0$ ),  $S_0^2$  coincides with the Poincaré sphere  $\Sigma_1^2$ . It is important to visualize the asymmetric depolarization, as  $n$  is increased, through the change in symmetry of the surface  $S_2^{P(n)}$  passing from a sphere ( $n = 0$ ) to a prolate ellipsoid ( $n > 0$ ). We thus have

drawn the surface  $S_2^{P(n)}$  in Fig. 4.1.15, where we present the result for different values of  $n$ .

An attempt has been made to express any physically realizable Mueller matrix as a sum of a nondepolarizing part  $\mathbf{M}_J$  and a pure depolarizing part proportional to  $\mathbf{D}$ . This decomposition is useful because, for any incident pure state of polarization, the decrease of the degree of polarization is solely caused by the depolarizing part, since light is still completely polarized after passing through the Mueller-Jones component.

#### 4.1.6. DIGRESSION: EXTENDED MATRIX FORMALISMS

The methods discussed thus far are not entirely adequate to deal with the polarization behavior of electromagnetic waves at interfaces and in thin films, or when the incident angle of light is arbitrary. We shall now briefly outline some features of other matrix methods that have been used over the years to describe the propagation of polarized light through optical media.

1. In the Berreman formalism, Maxwell's equations are used to construct a  $4 \times 4$  matrix whose eigenvectors are column vectors containing two components of the electric field and two components of the magnetic field that are parallel to an interface [48]. Within this approach, any reflection and refraction problem becomes an eigenvalue-eigenvector problem.

2. The  $T$ -matrix method (also called the extended boundary condition method) was introduced by Waterman [49]. It allows for exact scattering computations for particles having arbitrary shapes and optical properties [50]. Basically, in this matrix formalism, the electric fields are expanded in vector spherical harmonic functions. The electric field vector, in the form of a plane wave with unit amplitude, that is incident on a particle may be written in the form

$$\mathbf{E}_i(q\mathbf{R}) = \sum_{l=1}^{\infty} D_l [a_l \mathbf{M}_l^1(q\mathbf{R}) + b_l \mathbf{N}_l^1(q\mathbf{R})] \quad (4.1.168a)$$

where  $a_l$  and  $b_l$  are the expansion coefficients of the incident plane wave that have simple analytic expressions,  $D_l$  is a normalization coefficient,  $E_o$  is the amplitude,  $\mathbf{M}_l^1$  and  $\mathbf{N}_l^1$  are vector spherical harmonics of the first kind (they have a Bessel function radial dependence and are regular at the origin),  $q = 2\pi/\lambda$ ,  $\lambda$  is the free-space wavelength, and  $\mathbf{R}$  is a position vector. The scattered field vector by a particle is written

$$\mathbf{E}_o(q\mathbf{R}) = \sum_{l=1}^{\infty} D_l [f_l \mathbf{M}_l^3(q\mathbf{R}) + g_l \mathbf{N}_l^3(q\mathbf{R})] \quad (4.1.168b)$$

where  $f_l$  and  $g_l$  denote the expansion coefficients of the scattered field and are

initially unknown, and  $\mathbf{M}_l^3$  and  $\mathbf{N}_l^3$  are vector spherical harmonics of the third kind (they have a Hankel function radial dependence and vanish at infinity). By virtue of the linearity of Maxwell's equations, the incident and the scattered electric fields are related by a linear transformation

$$\begin{bmatrix} f_l \\ g_l \end{bmatrix} = \mathbf{T} \begin{bmatrix} a_l \\ b_l \end{bmatrix} \quad (4.1.168c)$$

where the elements of the  $\mathbf{T}$  matrix contain integrals over the surface of the scatterer and depend on the particle shape and optical properties and the orientation of the particle with respect to the coordinate system [50]. If the  $\mathbf{T}$  matrix for a given scattering configuration is known, Eqs. (4.1.168a,c) can be used to compute the scattered field. The basis functions are selected to give well-conditioned matrices and to permit analytic evaluation of the integrals involved. The elements of the Jones matrix can be expressed in terms of  $f_l$  and  $g_l$ , which, in turn permits computation of the Mueller matrix of the scattering.

3. In 1978 van Weert discussed a general  $2 \times 2$  matrix treatment for imaging systems in which polarization is incorporated. The Jones matrix formalism is, in a special case, incorporated in this analysis [51].

4. Although the normal incidence and paraxial assumptions are sufficient for many purposes, it may be necessary to develop the matrix formalism for treating the transmission of off-axis light through optical systems. The  $2 \times 2$  matrix method of Yeh can be applied to analyze this problem for sheet polarizers and birefringent crystal plates [52]. This method accounts for the effects of refraction and multiple reflections between interfaces in an arbitrary anisotropic layered medium. Yeh has shown that his method can be employed to explain the leakage of off-axis light through a pair of crossed ideal polarizers. Gu and Yeh have generalized this method to cover all dielectric media including uniaxial and biaxial crystals and also gyrotropic materials that exhibit optical rotation and Faraday rotation [53].

We now possess all the tools needed to analyze problems involving the propagation of waves through a scattering medium and the interaction of light with arbitrary optical element.

## ■ SECTION 4.2

---

# Polarization Effects at Dielectric Interfaces

Thus far, and especially in Part 2, we have been concerned with the problem of the transmission and reflection of a plane wave at a plane interface between two isotropic, homogeneous, dielectric media by Fresnel's equations. These equations govern the distribution of energy of the incident light into the refracted and reflected waves. In fact, we assumed that the incident wave was linearly polarized (parallel and perpendicularly polarized) and we considered only nondepolarizing media. For analysis of optical systems containing anisotropic materials, polarizers, depolarizers, or other elements, a formalism that account for depolarization is required. The question now quite naturally arises as to whether it might be possible to deal with any partially polarized state. A more general partially polarized wave treatment must be attempted. With this objective in mind, we present a complete determination of the polarization properties based on the Mueller matrix elements rather than investigating the p- (parallel) and s- [perpendicular (*senkrecht*)] polarized intensities alone as is conventionally done. Thus, in this chapter, we shall examine the general problem of the linear transformation of the Stokes vector by transmission and reflection at interfaces using the Mueller matrix description. In the analysis that follows, we will again consider a flat interface separating two homogeneous, isotropic, dielectric media. It should be noticed that this analysis is inadequate when the interface has significant roughness. This comes, in part, from the fact that the characterization of rough surfaces is not easy. Moreover, this limitation is also due to the difficulty in applying boundary conditions at these surfaces. We will assume that the interface is spatially uniform in its polarization modification effect.

The order of presentation goes as follows. Section 4.2.1 deals with the Mueller matrix formulation of Fresnel's equations. Section 4.2.2 briefly summarizes the basis principles of ellipsometry. In the analysis to follow we remark that all matrices are defined in the usual way for the linear polarization basis.

#### 4.2.1. MUELLER MATRIX FORMULATION OF FRESNEL'S EQUATIONS

In this section, we present a quantitative and illustrative Mueller matrix formulation of Fresnel's equations. To derive the reflection and transmission Mueller matrices, we need to use some basic results from the electromagnetic analysis of the Fresnel reflection and refraction at the static and plane interface between a dielectric medium and vacuum when a quasimonochromatic partially polarized plane wavefield in vacuum is incident on a dielectric interface. We once again consider the experimental arrangement depicted in Fig. 2.2. From Part 2, we already know that all optical interfaces display some polarization effects when used with light at nonnormal incidence. The incident wave is, in general, partially reflected and partially transmitted according to Fresnel's laws.

The matrices  $\mathbf{R}$  and  $\mathbf{T}$  have only four unique nonzero elements  $r_{ij}(\theta_i, \theta_t)$  and  $t_{ij}(\theta_i, \theta_t)$ . The following expressions for the reflection and transmission Mueller matrices for the dielectric interface displayed in Fig. 2.2 are written respectively [54]

$$\mathbf{R}(\theta_i, \theta_t) = \begin{bmatrix} a + b & a - b & 0 & 0 \\ a - b & a + b & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & c \end{bmatrix} \quad (4.2.1)$$

and

$$\mathbf{T}(\theta_i, \theta_t) = \begin{bmatrix} a' + b' & a' - b' & 0 & 0 \\ a' - b' & a' + b' & 0 & 0 \\ 0 & 0 & c' & 0 \\ 0 & 0 & 0 & c' \end{bmatrix} \quad (4.2.2)$$

where we have set

$$\begin{aligned} a &= \frac{1}{2} \left| \frac{\tan(\theta_i - \theta_t)}{\tan(\theta_i + \theta_t)} \right|^2, & b &= \frac{1}{2} \left| \frac{\sin(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t)} \right|^2, & c &= \frac{\sin(\theta_i - \theta_t) \tan(\theta_i - \theta_t)}{\sin(\theta_i + \theta_t) \tan(\theta_i + \theta_t)}, \\ a' &= \frac{1}{2} \left| \frac{2 \sin(\theta_t) \cos(\theta_i)}{\sin(\theta_i + \theta_t) \cos(\theta_i - \theta_t)} \right|^2, & b' &= \frac{1}{2} \left| \frac{2 \sin(\theta_t) \cos(\theta_i)}{\sin(\theta_i + \theta_t)} \right|^2, \\ c' &= \frac{4 \sin^2(\theta_t) \cos^2(\theta_i)}{\sin^2(\theta_i + \theta_t) \cos(\theta_i - \theta_t)} \end{aligned}$$

and where  $\theta_i$  and  $\theta_t$  refer to the incident and transmitted angles, respectively.

The preceding equations are sufficiently interesting to deserve some comments. First it is important to appreciate that Eqs. (4.2.1) and (4.2.2) are similar to the Mueller matrix of a partial linear polarizer, namely, Eq. (4.1.125). Consequently the interface behaves as a partial polarizer. We will utilize this result shortly in our discussion of the pile-of-plates polarizer. Unpolarized light reflected off of a plane surface such as that displayed in Fig. 2.2 becomes partially linearly polarized

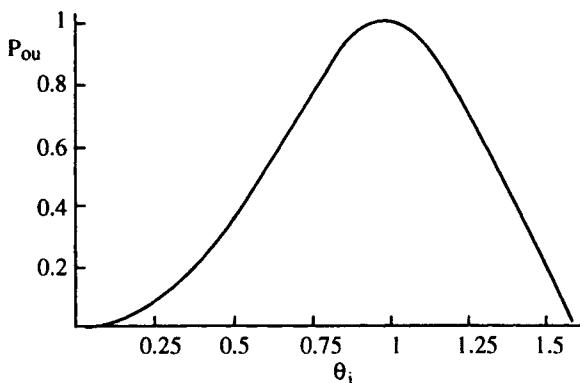
$$\langle S_0 \rangle \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow \langle S_0 \rangle \begin{bmatrix} a+b \\ a-b \\ 0 \\ 0 \end{bmatrix} \quad (4.2.3)$$

with a degree of polarization

$$P_{ou} = \frac{|a-b|}{a+b} \quad (4.2.4)$$

which is generally less than one. To examine the effect more closely, we have plotted  $P_{ou}$  in Fig. 4.2.1 as a function of  $\theta_i$ . Increasing  $\theta_i$  increases  $P_{ou}$  until it reaches a maximum for which  $P_{ou} = 1$ , corresponding to the Brewster angle, Eq. (2.1.23a). Past this point  $P_{ou}$  decreases to zero at  $\theta_i = \pi/2$ .

Before we proceed further, then, we must discuss the case of normal incidence of the light beam,  $\theta_i \equiv 0 = \theta_t$ . As Eqs. (4.2.1) and (4.2.2) shows, we



**Figure 4.2.1.** Degree of polarization as a function of the incident angle  $\theta_i$  for incident unpolarized light. Light is incident on an air/crown glass interface ( $n = 1.5$ ).

obtain

$$\mathbf{R}(0, 0) = \left(\frac{n-1}{n+1}\right)^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \quad (4.2.5a)$$

and

$$\mathbf{T}(0, 0) = n \left(\frac{2}{n+1}\right)^2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.2.5b)$$

where  $n$  is the ratio of the refractive index of the transmitted medium to that of the incident medium. Normal reflection from a flat surface preserves the direction of linear polarization but reverses the handedness (helicity) of circular polarization.

To close out this section, we shall now attempt to illustrate the use of these formulas for the special situation corresponding to the Brewster angle. Collett has shown that  $\mathbf{R}$  and  $\mathbf{T}$  become [54]

$$\mathbf{R}(\theta_B) = \frac{1}{2} \cos^2(2\theta_B) \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.2.6a)$$

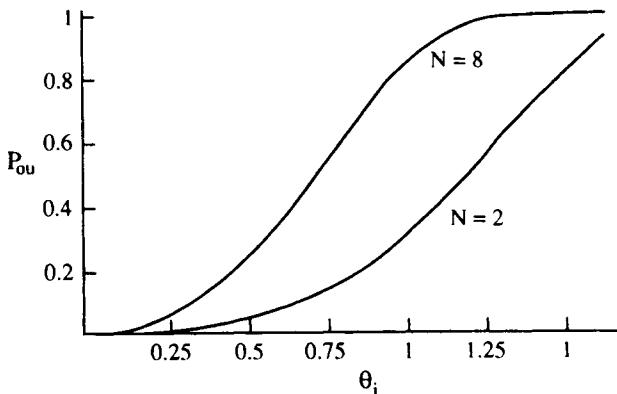
and

$$\mathbf{T}(\theta_B) = \frac{1}{2} \begin{bmatrix} \sin^2(2\theta_B) + 1 & \sin^2(2\theta_B) - 1 & 0 & 0 \\ \sin^2(2\theta_B) - 1 & \sin^2(2\theta_B) + 1 & 0 & 0 \\ 0 & 0 & 2 \sin(2\theta_B) & 0 \\ 0 & 0 & 0 & 2 \sin(2\theta_B) \end{bmatrix} \quad (4.2.6b)$$

where  $\theta_B$  is as given by Eq. (2.1.23a).

#### 4.2.2. THE PILE-OF-PLATES POLARIZER

In view of the preceding discussion, it is interesting to consider now the Mueller matrix of a dielectric plate. We assume that the plate consists of two plane interfaces without multiple internal reflections. Under these circumstan-



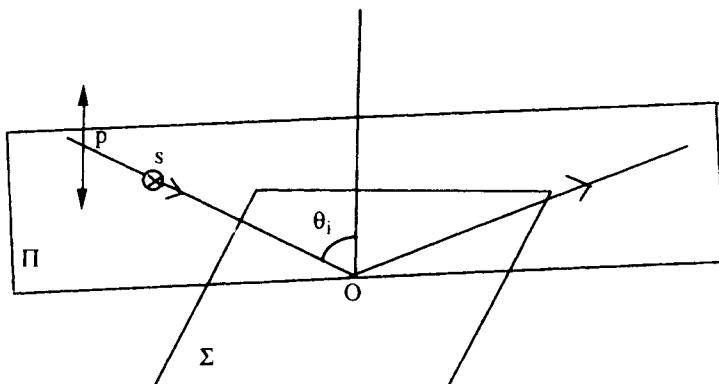
**Figure 4.2.2.** Degree of polarization as a function of the incident angle  $\theta_i$  for incident unpolarized light. Light is incident on a pile of  $N$  glass plates ( $n = 1.5$ ).

ces, the Mueller matrix of the plate is  $T^2(\theta_i, \theta_i)$ . An interesting problem is the extension of the analysis to cover the case of transmission through  $N$  parallel plates, where one has to deal with Eq. (4.1.90). We can utilize Eqs. (4.2.2) and (4.1.90) to evaluate the degree of polarization of the wave emerging from  $N$  parallel identical dielectric plates for an incident unpolarized light. The result is displayed in Fig. 4.2.2 as a function of the incident angle  $\theta_i$ .

As  $N$  increases, we see that the degree of polarization  $P_{ou}$  increases for a given incident angle. In the limit of very small angles  $\theta_i$ , that is, normal incidence, we remark that  $P_{ou} \cong 0$  regardless of the number of dielectric plates. Reference is made to Collett for the full details [54].

#### 4.2.3. PRINCIPLE OF ELLIPSOMETRY

An abbreviated introduction to ellipsometry will be the subject of this section. Ellipsometry is a technique for an *in situ* nondestructive characterization of surface (interfacial) phenomena utilizing the change in the state of polarization of a lightwave probe. The theory and practice of ellipsometry have been standardized for many years. There are many applications based on the retrieval of optical material properties from ellipsometric data. For example, in semiconductor technology, it is important to estimate the depth profile of the concentration of impurities caused by ion implantation or oxidization. Since the refractive index and the optical absorption coefficient are sensitive to electronic properties, many of these effects are correlated with the optical depth profile, which can be used for material characterization. The standard reference in the ellipsometry community is Azzam and Bashara [24]. Here we will examine the special case of reflection from an isotropic material. For optically isotropic surfaces, ellipsometry can be carried out only at oblique incidence



**Figure 4.2.3.** Definition of the plane of incidence at a point on a plane surface. The  $p$  wave is polarized so that its electric vector is within the plane  $\Pi$  of incidence. The  $s$  wave is polarized perpendicular to the plane of incidence. Intermediate linear polarization angles, such as  $+45^\circ$ , are measured counterclockwise from  $p$ , as viewed opposite to the direction of propagation.

since normal incidence reflection from such surfaces produces no change of polarization. This is not so for optically anisotropic surfaces, where a change of polarization can in general take place on normal incidence. The reader interested in the case of anisotropic materials, such as natural anisotropy for a cleaved crystal or induced by epitaxial deposition on a substrate, may wish to consult Refs. 55 and 56, for example.

Here we briefly discuss the Stokes–Mueller formalism for specifying the polarization dependence of the reflection of light from a plane interface. We assume a flat isotropic surface with reference to Fig. 4.2.3.

The changes in the polarization state of light reflected from the surface can be fully characterized by two parameters: the relative phase change  $\Delta$  and the relative change in amplitude  $\tan(\psi)$  between the incident and reflected components of light polarized parallel (resp.  $p$ ) and perpendicular (resp.  $s$ ) to the plane of incidence. These two parameters define the complex number  $Z \equiv \tan(\psi) \exp(i\Delta)$ , which represents the ratio of the  $p$  and  $s$  complex amplitude reflection coefficients  $r_p/r_s$ . From notations of Section 2.1.3, we have  $\tan(\psi) = |r_p|/|r_s|$  and  $\Delta = \phi_{r_p} - \phi_{r_s}$  indicating that the two parameters characterize the phase and amplitude changes of the electric vibration on reflection and transmission. Now, substitution of Eqs. (2.1.21a) and (2.1.21b) into the definition of  $Z$  and making use of the Descartes–Snel law leads to the important expression

$$\frac{Z}{(1 + Z^2)} = \frac{\tan(\theta_i^2) + 1 - n^2}{4 \sin(\theta_i^2)} \quad (4.2.7)$$

where  $n \equiv n_2/n_1$  is the notation for the relative index of refraction and  $\theta_i$  is the angle of incidence. Equation (4.2.7) allows us to calculate the complex index of refraction of the reflecting bulk material from a single measurement of the two parameters  $\Delta$  and  $\psi$ , provided  $n$  and  $\theta_i$  are known.

The Jones matrix describing the transformation from incident wave to reflected wave can be expressed as

$$\mathbf{J}_e(\psi, \Delta) = \begin{bmatrix} (1 - \cos(2\psi))^{1/2} & 0 \\ 0 & \frac{\sin(2\psi) \exp(i\Delta)}{(1 - \cos(2\psi))^{1/2}} \end{bmatrix} \quad (4.2.8)$$

Note that the matrix  $\mathbf{J}_e(\psi, \Delta)$  contains off-diagonal terms when the interface is anisotropic [24].

When the surface is flat, that is, when the Fresnel equations are valid, the Mueller matrix for the surface is given by [24]

$$\mathbf{M}_e(\psi, \Delta) = \begin{bmatrix} 1 & -\cos(2\psi) & 0 & 0 \\ -\cos(2\psi) & 1 & 0 & 0 \\ 0 & 0 & \sin(2\psi) \cos(\Delta) & \sin(2\psi) \sin(\Delta) \\ 0 & 0 & -\sin(2\psi) \sin(\Delta) & \sin(2\psi) \cos(\Delta) \end{bmatrix} \quad (4.2.9)$$

An important observation is that  $\mathbf{E}(\psi, \Delta)$  can be decomposed as a commutative product of a compensator  $\mathbf{C}(-\Delta)$ , Eq. (4.1.107), and a partial linear polarizer  $\mathbf{P}(\psi)$ , Eq. (4.1.119), whose maximum and minimum gains are respectively given by  $g_{\max} = \sin^2(\psi)$  and  $g_{\min} = \cos^2(\psi)$ . Inspection of Eq. (4.2.9) leads to the important conclusion that no depolarization and no cross-polarization occur that for a flat surface, that is, pure p state

$$\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

and s state

$$\begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix}$$

are not mapped to themselves. As an illustrative example, we consider an

incident right-handed circular polarized wave. Equation (4.2.9) yields the output Stokes vector

$$\mathbf{S}_o = \begin{bmatrix} 1 \\ -\cos(2\psi) \\ \sin(2\psi) \sin(\Delta) \\ \sin(2\psi) \cos(\Delta) \end{bmatrix}$$

Hence, by measuring the output Stokes parameters, one can determine the two parameters  $\psi$  and  $\Delta$ :

$$\tan(\Delta) = \frac{\langle S_{o2} \rangle}{\langle S_{o3} \rangle} \quad (4.2.10a)$$

$$\tan(2\psi) = \frac{(\langle S_{o2} \rangle^2 + \langle S_{o3} \rangle^2)^{1/2}}{\langle S_{o1} \rangle} \quad (4.2.10b)$$

When the surface is sufficiently rough, that is, having steep slopes and structures comparable in size to the illuminating wavelength, depolarization effects may be predominant and require that other parameters than  $\Delta$  and  $\psi$  be introduced to characterize its optical properties. Such difficulties are extensively surveyed in the literature.

In summary, in this section we have described methods for specifying the polarization effects at dielectric interfaces and we have illustrated these results in several simple cases. In the next section, we address the important question of the influence of symmetry transformations on polarization matrices.

---

## SECTION 4.3

---

# Polarized Light and Symmetry Transformations

Fundamental physical laws are subject to a variety of symmetry transformations. In particular, we can get at the properties of a system without completely solving all the equations that describe the system. These symmetry properties have played an important role in recent investigations of elastic scattering of light by ensemble of particles. The useful symmetries of concern are of different kinds (e.g., time reversal, spatial inversion). If we apply spatial symmetry operations (e.g. rotations, rotation reflections), it leads to special features of the polarization matrices representing the effect of optical devices that transform both the Jones vector and the Stokes vector. On the basis of the invariance under these symmetry transformations in an appropriate basis, the polarization matrix elements are related by a set of equations that restrict the number of independent parameters. Work on symmetry has been hampered on several fronts. Reference is made to the pioneering researches of Perrin [57], Krishnan [58], and Chandrasekhar [59] as well as to van de Hulst [60], for an overview. Much less attention has been given to understanding the consequences of time reversal symmetry, that is, when the incident and scattered beams of light are interchanged or when their directions of propagation are reversed. It is interesting to note that in geometric optics, the reversibility of ray propagation originates from Fermat's principle.

The primary purpose of this section is to study symmetry transformations on polarization matrices. An additional purpose is to provide the information needed to characterize the quantitative effects of perfect exactly known scattering systems such as small spheres on polarization of light. It is noteworthy that such perfect systems can be used to evaluate methods of measurement of polarization matrices because their interaction with partially polarized light can be exactly predicted from first principles.

The order of presentation is as follows. Section 4.3.1 is devoted to the presentation of symmetry transformations for Jones and Mueller matrices. In particular, we consider the scattering of light by an ensemble of particles having different kinds of symmetries. In particular, we shall see that symmetry arguments imply that the number of independent Mueller scattering matrix

elements is actually less than 16. Section 4.3.2 considers time-reversal invariance and reciprocity. Here the purpose is to present an analytic calculation showing that time-reversal symmetry implies the existence of specific restrictions on polarization matrices.

### 4.3.1. SPATIAL SYMMETRY RELATIONS FOR A FAR-FIELD SCATTERING

The scattering of light by particulate media is an old field. The fundamentals of single-scattering theory were established in the scientific literature of the late nineteenth and early twentieth centuries. Particularly noteworthy early contributions were made independently by Lord Rayleigh, Lorenz in 1890 and Mie in 1908. The most important and influential works on the polarization effects on scattering were perhaps done by Perrin, Krishnan, and Chandrasekhar [57–59]. For a detailed historical review of the general subject of elastic scattering of an electromagnetic wave, the reader may refer to Part 1. In more recent times, important developments are found in the classic text by van de Hulst [60]. The development of this section lean heavily on this original work.

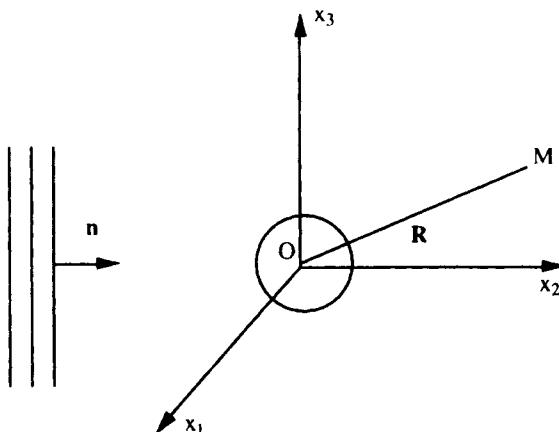
Here we shall only give some of the major categories of polarization matrices that describe scattering of light by particles. The objective is to help take advantage of polarization properties by utilizing the information contained in the polarization matrices. In most applications what matters are the features of particular cases, rather than a systematic general theory, so few optical physicists need master the full analysis of the rigorous, first-principles scattering theory based on electrodynamics such as given by Newton [61]. This has useful consequences in a wide variety of situations in optics, such as remote sensing techniques, lidar and radar polarimetry, and to characterize the shapes of scattering particles. There has been extensive research on elastic light scattering by ensembles of particles for a variety of applications, such as scattering by interstellar dust in astronomy, and optical properties of hydrosols in ocean optics.

The passage of light through an optical medium is, in general, accompanied by the removal of a fraction of the energy from the incident beam. This fraction of electromagnetic radiation may be partly absorbed within the medium and may become partly scattered, that is, redistributed in other directions of space. We shall assume media that are optically thin; thus the situation can be treated by the single-scattering approximation. In the opposite situation, the wave undergoes multiple-scattering events, some of which redirect the wave back to the forward direction, thus increasing the transmitted fraction of incident flux. This topic will be discussed in Section 4.4. The derivation here focuses on the description of electromagnetic scattering for a single particle. The properties of an ensemble are then obtained by an averaging over a large number of configurations. In this section we will consider ensembles of monodisperse

particles that are randomly oriented with no preferred (left- or right-) handedness. A monodisperse ensemble is one that has all its particles identical, in contradistinction to a polydisperse ensemble. Then all the statistical averages are mirror-symmetric.

In the elastic scattering case of interest here, the mean frequency of the scattered radiation remains the same as the mean frequency of the incident radiation. Any scattering medium is composed of discrete scatterers. In most practical applications we shall be interested in the scattered light at large distances. At such distances we are in the radiation zone where the fields have no radial components, that is, the field is closely approximated by a plane wave. In practice, it is well established that the plane-wave approximation is valid for scattering from particles that are much narrower than the incident beam.

The general problem of elastic scattering of an electromagnetic wave by an isotropic, nonmagnetic material with no sources is perhaps illustrated best by considering the simplest case of a sphere of arbitrary refractive index. We consider an incident plane wave propagating in the direction of a unit vector  $\mathbf{n}$ . The wave is elastically scattered by a single dielectric spherical particle (of radius  $a$ ) located at the origin of the coordinate system. The geometry of the problem is defined in terms of a local right-handed Cartesian coordinate system  $(x_1, x_2, x_3)$  with origin inside the particle, as indicated in Fig. 4.3.1. The perspective view in this figure identifies the direction of light propagation, which is specified by  $\theta$  and  $\phi$ , where  $\theta$  is the polar angle and  $\phi$  is the azimuthal



**Figure 4.3.1.** Notation related to the scattering geometry discussed in the text. An incident plane wave is scattered by a dielectric sphere located at the origin of the coordinate system. The term  $\mathbf{n}$  is a unit vector in the direction of propagation of the incident field. Point  $M$  is a typical observation point in the far zone. The quantity  $R$  is the distance from the origin to the distant observation point  $M$ .

angle, in the coordinate system presented. The scattered wave is monitored by a light detector in the direction  $\mathbf{n}'(\theta, \phi)$  at a distance  $R$  in the far field. We wish to determine the appropriate expression for the Jones and Mueller polarization matrices in the far zone:  $|\mathbf{qR}| \gg 1$ .

In the process of scattering, the incident plane wave is transformed to an outgoing spherical wave. Because in most experimental situations the polarization matrices are deduced from the far-field data, the wave can be treated as planar. The linearity of the Maxwell equations allows one to write down the scattering Jones matrix in the far field by the following equation

$$\begin{aligned} \mathbf{E}_o(\mathbf{n}') &= \lim_{qR \rightarrow \infty} \frac{\exp(iqR)}{iqR} \mathbf{J}(\theta, \phi) \mathbf{E}_i(\mathbf{n}) \\ &= \lim_{qR \rightarrow \infty} \frac{\exp(iqR)}{iqR} \begin{bmatrix} J_{11}(\theta, \phi) & J_{12}(\theta, \phi) \\ J_{21}(\theta, \phi) & J_{22}(\theta, \phi) \end{bmatrix} \mathbf{E}_i(\mathbf{n}) \end{aligned} \quad (4.3.1)$$

where  $q$  is the wavenumber of the host medium. Since  $q$  and  $R$  are not properties of the scatterer, they may be excluded from the scattering Jones matrix  $\mathbf{J}(\theta, \phi)$ . The factor  $1/iqR$  makes the matrix  $\mathbf{J}(\theta, \phi)$  dimensionless. Since the scattering is elastic, the wavenumber of the outgoing wave is the same as that of the incident wave. This matrix, also termed the *scattering amplitude matrix*, describes how the components of a plane wave  $\mathbf{E}_i$  incident on the spherical particle in the direction are modified by scattering. It is a function of  $\mathbf{n}$  and  $\mathbf{n}'$  and is expressed in the particle coordinate system. It also depends on the size, morphology, and composition of the scattering particle as well as the particle's orientation with respect to the coordinate system. For a collection of identical randomly oriented particles, only the dependence on the scattering angle  $\theta$  needs be considered; the azimuthal angle  $\phi$  dependence can be set to zero. Transformation of the Jones matrix from the representation in the particle coordinate to the medium may be done by using Euler angle transformation.

There exist several methods for the explicit calculation of such a matrix. The analytic expressions of the amplitude  $J_{kl}$ 's values for a spherical particle or small spheroid can be found in most references. To illustrate this point, consider the case of scattering by small particles. The scattering matrix by an isotropic particle with isotropic polarizability  $\alpha$  may be written as

$$\mathbf{J}(\theta) = iq^3\alpha \begin{bmatrix} \cos(\theta) & 0 \\ 0 & 1 \end{bmatrix} \quad (4.3.2)$$

An explicit evaluation of the gain for an incident unpolarized wave can be accomplished by using Eqs. (4.1.24) and (4.3.2), yielding the Rayleigh law  $g_u \sim [1 + \cos^2(\theta)]/2$ .

Here we shall be concerned with media composed not of one single particle but of many particles incoherently scattering the wave. With knowledge of the

Jones scattering matrix by an arbitrary particle, we then introduce statistical information on the number and distribution of particles. If all the particles were identical and oriented in the same direction,  $\mathbf{J}(\mathbf{n}, \mathbf{n}')$  would be a deterministic matrix. However, in real situations, the particles are polydispersed and have random orientations. Consequently, the Jones matrix should be ensemble-averaged of these quantities. The reader is referred to van de Hulst for details of the form of  $\mathbf{J}$  if certain assumptions about the distribution of orientations are made [60].

Thus far the analysis has been confined to optical media that are characterized by Jones matrices. An extension of the analysis to include Mueller matrices is of interest. Perrin was the first to give a detailed study of the number of independent parameters (among the 16 matrix elements of  $\mathbf{M}$ ) that are necessary for specifying the polarization characteristics of light scattered by an arbitrary medium. Following the analysis made by Perrin, the forward (or backward) axial scattering by a symmetric medium (e.g., identical particles having spherical symmetry) involves only three coefficients; moreover, the Mueller matrix is diagonal. We note in passing that further refinements for a variety of symmetry properties of the  $\mathbf{M}$  matrices describing the reflection and transmission of polarized radiation by a slab of randomly oriented particles have been treated by van de Hulst and later by Hovenier. In a statistically isotropic three-dimensional medium, there is full  $C_{\infty V}$  symmetry around the propagation direction. In two-dimensions, the symmetry is reduced to  $C_{2V}$ . The Mueller matrix is not diagonal but block diagonal in this case. In recent studies of elastic light scattering by ensembles of particles, a number of authors found that certain elements are sensitive to structural information of particles, including size and shape information as well as information on whether the particles are randomly oriented; for instance,  $m_{34}$  has a specific structured angular distribution and can serve as a diagnostic for distinguishing among bacterial species and other well-structured cells [67–69].

The scattering process can be described as a linear transformation of the Stokes vector of the incident wave into the Stokes vector of the scattered wave through the Mueller matrix:

$$\mathbf{S}_o = \frac{1}{(kR)^2} \mathbf{N}(\theta, \phi) \mathbf{S}_i = \frac{1}{(kR)^2} \begin{bmatrix} n_{00}(\theta, \phi) & n_{01}(\theta, \phi) & n_{02}(\theta, \phi) & n_{03}(\theta, \phi) \\ n_{10}(\theta, \phi) & n_{11}(\theta, \phi) & n_{12}(\theta, \phi) & n_{13}(\theta, \phi) \\ n_{20}(\theta, \phi) & n_{21}(\theta, \phi) & n_{22}(\theta, \phi) & n_{23}(\theta, \phi) \\ n_{30}(\theta, \phi) & n_{31}(\theta, \phi) & n_{32}(\theta, \phi) & n_{33}(\theta, \phi) \end{bmatrix} \mathbf{S}_i \quad (4.3.3)$$

The 16 elements of the Mueller matrix of an arbitrary scatterer contain all the detailed information about the structural (size, shape) and optical properties (refractive index) in an elastic scattering process. Each of the 16 elements of the Mueller matrix depends on the scattering angle  $\theta$  and the azimuthal angle  $\phi$ .

A usual medium in light scattering studies is a collection of particles in which all possible orientations of a single particle are equally probable (e.g., a dilute suspension of particles). If multiple-scattering effects among the particles can be neglected, the Mueller scattering matrix of the medium will correspond to the averages values for all possible orientations of a single particle. To evaluate the Mueller matrices of isotropic ensemble of uncorrelated scatterers, we must perform the following operations:

1. Solve Maxwell's equations for the radiative field.
2. Convert the field to the Mueller matrix observables of a single scatterer in a particular orientation.
3. Average the observables over all orientations of the scatterer.

Further, we must determine what simplifications of the Mueller matrices we can expect if certain symmetry assumptions in a statistical sense are made? It is useful to consider some simple examples:

1. For a randomly oriented collection of particles, the number of nonzero independent matrix elements reduces to 10 if we exclude particles with optical activity. This results in the following type of Mueller matrix, which we designate as  $\mathbf{M}_1$ :

$$\mathbf{M}_1 = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{01} & m_{11} & m_{12} & m_{13} \\ -m_{02} & -m_{12} & m_{22} & m_{23} \\ m_{03} & m_{13} & -m_{23} & m_{33} \end{bmatrix} \quad (4.3.4)$$

2. In an assembly of randomly oriented particles, each of which has a plane of symmetry (e.g., finite cylinders), if the particles and their mirror images are in equal numbers and with random orientation, this latter type of Mueller matrix reduces to  $\mathbf{M}_2$ :

$$\mathbf{M}_2 = \begin{bmatrix} m_{00} & m_{01} & 0 & 0 \\ m_{01} & m_{11} & 0 & 0 \\ 0 & 0 & m_{22} & m_{23} \\ 0 & 0 & -m_{23} & m_{33} \end{bmatrix} \quad (4.3.5)$$

3. For the special case of spherical particles, symmetry arguments reduce the Mueller matrix elements to four independent matrix elements. The Mueller

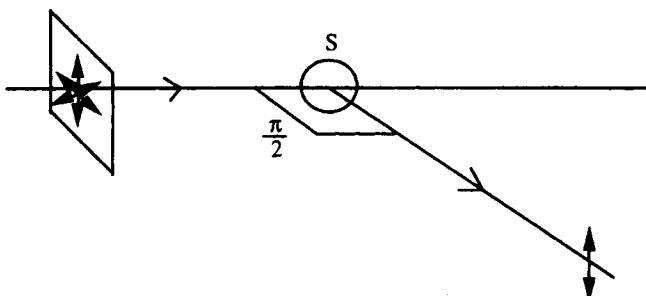
matrix  $\mathbf{M}_3$  has the form

$$\mathbf{M}_3 = \begin{bmatrix} m_{00} & m_{01} & 0 & 0 \\ m_{01} & m_{00} & 0 & 0 \\ 0 & 0 & m_{22} & m_{23} \\ 0 & 0 & -m_{23} & m_{22} \end{bmatrix} \quad (4.3.6)$$

Clearly these relations can serve as signatures of the scatterer and may be used to identify the scatterer to a certain extent. We note that if  $m_{11} \neq m_{00}$ , or  $m_{33} \neq m_{22}$ , this reflects the fact that the scattering medium is composed of anisotropic particles. The scattering Mueller matrix can carry the imprint of the topology of the scatterers, such as an indicator of asphericity is the degree to which the parameter  $1 - (m_{11}/m_{00})$  diverges from zero.

A scatterer can be completely described by measuring all its matrix elements as a function of  $\theta$ . The Mueller scattering elements can be evaluated in closed form for spheres, spheroids, and infinitely long cylinders. For particles having shapes other than these ones, numerical solutions exist, such as the T-matrix method [50], or the volume integral equation formulation [62].

With these physical characteristics in mind, we will now briefly discuss a few general features of the Mueller matrix elements for Rayleigh spheres. The Mueller matrix for Rayleigh scattering is of the  $\mathbf{M}_3$  type with a simple angular dependence. The matrix elements are  $m_{00} = (1 + \mu^2)/2$ ,  $m_{01} = (\mu^2 - 1)/2$ ,  $m_{22} = \mu$ , and  $m_{23} = 0$ , where  $\mu = \cos(\theta)$  and  $\theta$  is the scattering angle. An example of the use of this result is deferred to the next section. An interesting property of a homogeneous spherical particle is that no change in the incident state of polarization occurs during scattering into backward and forward directions. Moreover, these elements are independent of particle size, shape, and optical properties as long as the particles are much smaller than the wavelength of the scattered radiation. For unpolarized incident wave, it follows that the intensity after a single scattering is given by the familiar



**Figure 4.3.2.** Rayleigh scattering.

function  $g_u = m_{00} = (1 + \mu^2)/2 = [1 + \cos^2(\theta)]/2$ , and we again obtain the well-known Rayleigh scattering law. At right angle to the direction of propagation, the reradiating wave is linearly vertically scattered (Fig. 4.3.2).

It is also worth noting that the Mueller matrix for perfectly aligned cylinders with their symmetry axis perpendicular to the scattering plane is of the  $\mathbf{M}_3$  type.

A special but important case deals with the scattering of an ensemble of randomly oriented identical spheroidal particles. The particles are sufficiently small compared to the wavelength of light, thus the Mueller matrix is of the  $\mathbf{M}_2$  type. A short derivation will be given here, but a more detailed treatment of this topic is to be found in solution of Asano and Sato [63]. The appropriate expressions for the Mueller matrix elements are quoted from Ref. 63:

$$m_{00} = A_1(1 + \mu^2) + \frac{2}{15}B^2 \quad (4.3.7a)$$

$$m_{01} = A_1(\mu^2 - 1) \quad (4.3.7b)$$

$$m_{11} = A_1(1 + \mu^2) \quad (4.3.7c)$$

$$m_{22} = 2A_1\mu \quad (4.3.7d)$$

$$m_{23} = 0 \quad (4.3.7e)$$

$$m_{33} = 2(A_1 - \frac{1}{15}B^2)\mu \quad (4.3.7f)$$

where  $A_1 = \frac{1}{2}(A^2 + \frac{2}{3}AB + \frac{2}{15}B^2)$ . For spheroidal particles,  $A$  and  $B$  may be expressed as

$$A = (ka)^3 \frac{(m^2 - 1)}{(1 + (m^2 - 1)L_b)3p^2} \quad (4.3.8a)$$

$$B = (ka)^3 \frac{(m^2 - 1)}{(1 + (m^2 - 1)L_a)3p^2} \quad (4.3.8b)$$

where  $m \equiv n_S/n_M$  is the relative refractive index,  $n_S$  is the refractive index of the particles, and  $n_M$  is the refractive index of the medium. The shape parameters  $L_i = a, b$  are given by

$$L_b\{p > 1\} = \frac{p^2}{2(p^2 - 1)} \left( 1 - \frac{1}{p(p^2 - 1)^{1/2}} \ln(p + (p^2 - 1)^{1/2}) \right) \quad (4.3.9a)$$

$$L_b\{p < 1\} = \frac{p^2}{2(p^2 - 1)} \left( \frac{1}{p(p^2 - 1)^{1/2}} \arccos(p) - 1 \right) \quad (4.3.9b)$$

where  $p \equiv a/b > 1$  for a prolate spheroid (rod) and  $p < 1$  for an oblate spheroid (disk) and

$$L_a + 2L_b = 1 \quad (4.3.10)$$

Note that for an isotropic spherical particle, that is, Rayleigh scattering, we have  $L_a = L_b = \frac{1}{3}$  and  $B \equiv 0$ . We observe from Eqs. (4.3.7a–f) that, when a plane wave is incident on a cloud of randomly oriented spheroidal scatterers, the scattered wave is not completely polarized even at a scattering angle of  $\pi/2$ , contrasting with the case of spheres. The scattered light is, in general, partially polarized.

We close this section with the following comment concerning the use of the coupled dipole approximation to the problem of evaluating Mueller scattering matrices. This technique, originally formulated by Purcell and Pennypacker [64] in 1973, has been also used to model light elastically scattered by particles of arbitrary shape and of size comparable to the wavelength. This method divides a dielectric particle, of moderate refractive index, into a number of identical elementary units located on the sites of a cubic lattice, which are assumed to behave as spherical dipolar oscillators. This assumption is invalid when the wavelength of light is of the same order of the size of the subunit. The estimation of the dipole moments is determined by writing the acting field at each site, involving the relation between the dipole moment and the electric field, and by use of the Clausius–Mossotti (or Lorentz–Lorenz) formula of the polarizability. The interactions between the dipoles are taken into account by determining self-consistently the field at a particular dipole due to the incident field as well as the fields induced by other oscillators. The scattered field is then obtained by summation of the fields radiated by all dipoles. Within this method, one is able to calculate all 16 elements of the Mueller matrix for an arbitrary-shaped particle, by performing the calculation for both parallel and perpendicular polarized incident electric fields to the scattering plane [69]. Several algorithms have been developed for the implementation of this model, and Mueller scattering matrix elements have been calculated by various authors for a number of geometries, including nonspherical particles [65], cubes [66], thin cylinders of finite length [67], infinite circular cylinders [68], spheroidal particles [63], and randomly oriented particles [69].

### 4.3.2. TIME-REVERSAL INVARIANCE AND RECIPROCITY

Here we wish to address the issue of reciprocity constraint on polarization transformation matrices, specifically, the behavior of the state of polarization in an optical system when the direction of propagation is reversed, that is, when the incident and scattered beams are interchanged. The principle of reciprocity (also termed the *Helmholtz reciprocity theorem*) in its general form states that the input and output of any linear passive optical system can be interchanged without altering the response of the optical system. An optical medium that satisfies the principle of reciprocity is termed a *reciprocal medium*; otherwise it is called a *nonreciprocal medium*. This reciprocity is a consequence of the time-reversal invariance requirement on Maxwell's equations.

It is useful to first examine, in Section 4.3.2.1, the consequences of time-reversal invariance. The method employed in Section 4.3.2.1 follows from

group theory. Next, Section 4.3.2.2 is devoted to the analysis of reciprocity relationships and their physical consequences.

### 4.3.2.1. Time Reversal Invariance

Our purpose here is to derive the form of polarization matrices. We begin by outlining the basic physical framework in terms of which the problem under investigation can be approached. Next we address the important question of precisely what this particular symmetry actually means operationally, and, in so doing, endeavor to explain, in physical terms, why a Jones matrix is defined within a unitary operator—a statement that is often quoted in the literature without any justification. Finally, an example is presented suggesting that this formalism is applicable in the characterization of polarization transformations by optical devices. Reference is made to Brosseau for the full details [70].

Let us now proceed with the physics of this problem. We consider an optical field in the form of narrowband plane waves of infinite extent that is incident on a deterministic and linear (nonmagnetic) optical system. We respectively denote by  $\mathbf{E}_i$  and  $\mathbf{E}_o$  the input and output Jones vectors of a nonimage-forming optical device characterized by a  $2 \times 2$  Jones complex-valued matrix  $\mathbf{J}$ . Our convention is that all quantities are defined in the linear polarization basis.

The basic point that I wish to stress follows the general theory of symmetry transformations; for a review, see Wigner [37]. Denoting by  $\mathbf{T}$  the operation of time reversal, we may represent this operation by the  $2 \times 2$  transformation matrix such that

$$\mathbf{T}\mathbf{E} = \mathbf{U}\mathbf{K}\mathbf{E} = \mathbf{U}\mathbf{E}^* \quad (4.3.11)$$

where  $\mathbf{U}$  is an element of the unitary group  $U(2)$  and  $\mathbf{K}$  is the operation of complex conjugation. This form of writing  $\mathbf{T}$  ensures that  $\mathbf{T}$  is an antiunitary and antilinear operator, which is expected from the very first principles. Now the constraint of time-reversal invariance on the Jones matrix is expressed by the commutation relation

$$\mathbf{T}\mathbf{J} = \mathbf{J}\mathbf{T} \quad (4.3.12)$$

By substituting Eq. (4.3.12) into Eq. (4.3.11), we obtain the analytic form of the time-reversal invariance condition:

$$\mathbf{J}^* = \mathbf{U}^{-1}\mathbf{J}\mathbf{U} \quad (4.3.13)$$

This result is sufficiently interesting to warrant some comments. First, it indicates that any Jones matrix is defined within a unitary operator, that is, determined up to an overall (nonmeasurable) phase. Second, it must be borne in mind that Eq. (4.3.13) is under the dependence of the particular unitary

representation chosen. What should also be pointed out is the involutorial nature of the time inversion. Therefore,  $\mathbf{T}$  must satisfy  $\mathbf{T}^2 = c\sigma_0$ , with  $c = \pm 1$ . Combined with Eq. (4.3.11), we can conclude that  $\mathbf{U}^* = c\mathbf{U}^+$  with  $c^2 = 1$ .

At this stage, it is appropriate to illustrate this parametrization of the Jones matrices by a couple of examples. In the first example, we take  $\mathbf{U}_1 = \sigma_0$ . Going back to Eq. (4.3.13), we find that the Jones matrix  $\mathbf{J}_1$  is real, that is, depending at most of four parameters. Examples of optical components that can be described by Jones matrices satisfying Eq. (4.3.13) include linear ideal polarizers and rotators. In that case, we have  $c = 1$ . Note that this operation is akin to phase conjugation. In the second example, we take  $\mathbf{U}_2 = -i\sigma_2$ , where

$$\sigma_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Thence Eq. (4.3.13) implies that

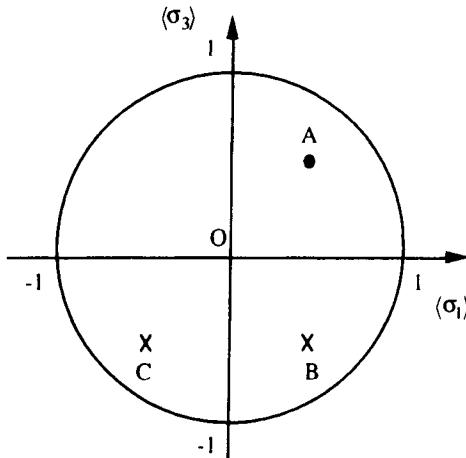
$$\mathbf{J}_2 = \begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix}$$

A typical example member of this class of Jones matrices is the compensator. Observe now that  $c = -1$ .

We turn next to another visualization of the preceding results that is based on the Poincaré space description of polarization states. In the light of the preceding results, we proceed to examine the constraints of time-reversal symmetry when it acts on the coherency matrix and on the Stokes vector. We first point out that the Stokes–Mueller description eliminates the absolute phase of the Jones formalism. Now returning to our previous example  $\mathbf{U}_1 = \sigma_0$ , we see that the transformation of the coherency matrix under the time reversal symmetry is simply a conjugation operation,  $\Phi \rightarrow \Phi^*$ . The corresponding Mueller matrix can be cast in the form

$$\mathbf{M}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

indicating that a Stokes vector, represented by point  $A$  in Poincaré space, is transformed into a Stokes vector, represented by point  $B$  (see Fig. 4.3.3). It appears that the matrix  $\mathbf{M}_1$  behaves like a circular switch, transforming a right-handed circular state into a left-handed circular state  $C \rightarrow \circlearrowleft$  but does not affect linear states of polarization.



**Figure 4.3.3.** Transformation of Stokes vectors by Mueller matrices  $\mathbf{M}_1$  and  $\mathbf{M}_2$  using the stereographic projection of the Poincaré sphere  $\Sigma_2^1$ . The circle  $\langle \sigma_1 \rangle^2 + \langle \sigma_3 \rangle^2 = 1$  is the circular boundary separating the physical (interior) and the nonphysical (exterior) states of polarization. Symbols ● and ✗ indicate that the points belong to the north and south hemispheres, respectively.

We now consider our second example  $\mathbf{U}_2 = -i\sigma_2$ . The transformation of the coherency matrix under the time-reversal symmetry is given by  $\Phi \rightarrow \det(\Phi)\Phi^{-1}$ . The Mueller matrix is now

$$\mathbf{M}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

to which corresponds the mapping  $A \rightarrow C$  in Poincaré space. Inspection of Fig. 4.3.4 clearly shows that the corresponding transformation rule for the Stokes vector is that the new Stokes vector is orthogonally polarized with respect to the old one. For instance, we have the following transformations of polarization states: linear horizontal state into linear vertical state  $\leftrightarrow \rightarrow \downarrow$ , linear  $+45^\circ$  state into linear  $-45^\circ$  state  $\nearrow \rightarrow \nwarrow$ , and right-handed circular state into a left-handed circular state  $C \rightarrow \odot$ .

The Pauli matrices have a special physical significance in polarization optics. First,  $\sigma_0$  characterizes an ideal, isotropic, nonabsorbing medium. The Jones matrix of a horizontal half-waveplate is proportional to  $\sigma_1$ , while the Jones matrix of a  $45^\circ$  half-waveplate is proportional to  $\sigma_2$ , and the Jones matrix of a right-handed circular half-waveplate is proportional to  $\sigma_3$ . From the explicit form of the Pauli matrices and substituting from Eq. (4.3.13), one

obtains the following commutation relations:

$$\mathbf{U}^{-1}\boldsymbol{\sigma}_1\mathbf{U} = \boldsymbol{\sigma}_1 \quad (4.3.14a)$$

$$\mathbf{U}^{-1}\boldsymbol{\sigma}_2\mathbf{U} = \boldsymbol{\sigma}_2 \quad (4.3.14b)$$

$$\mathbf{U}^{-1}\boldsymbol{\sigma}_3\mathbf{U} = -\boldsymbol{\sigma}_3 \quad (4.3.14c)$$

Practical implications of these equations are important because the Pauli matrices define a useful set of basis matrices; thus we may write any Jones matrix in the form  $\mathbf{J} = \sum_{i=0}^3 a_i \boldsymbol{\sigma}_i$ . We should note that the commutation relations (4.3.14a–c) induce specific forms of the Jones matrix in the linear polarization basis. These properties are summarized in Table 4.3.1. Another clear physical fact emerging from inspection of Table 4.3.1 is that the two unitary matrices  $\boldsymbol{\sigma}_0$  and  $-i\boldsymbol{\sigma}_2$  imply constraints on the Jones matrix that are very different. One must note that the four complex-valued coefficients  $a_i$  of the preceding expansion have definite physical meanings. For instance, the real (resp. imaginary) part of  $a_3$  characterizes the right circular birefringence (resp. dichroism) over left circular birefringence.

One can draw interesting applications in light of our analysis. As an important example of the application of such parameterization, consider the Faraday effect, in which linearly polarized light is rotated when it propagates through matter in the direction of an applied magnetic field, which breaks the time-reversal invariance. The Jones matrix for the linear polarization basis is  $\mathbf{R}(\theta)$  [i.e., Eq. (4.1.41)], where  $\theta = V \int_L H \cdot d\mathbf{l}$  is the Faraday rotation angle;  $V$  is the Verdet constant,  $H$  is the magnetic field, and  $L$  is the medium length. Equation (4.3.13) implies that

$$\mathbf{U} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

**TABLE 4.3.1. Transformation of Coherency Matrix, Commutation Relations [Eqs. (4.3.14a – c)] Satisfied, and Corresponding Form of Jones Matrix According to Choice of Unitary Matrix  $\mathbf{U}^a$**

$\mathbf{U}$	Transformation of Coherency Matrix	Commutation Relations Satisfied	$\mathbf{J}$
$\boldsymbol{\sigma}_0$	$\Phi \rightarrow \Phi^*$	(4.3.14a), (4.3.14b)	$\begin{bmatrix} a_0 + a_1 & a_2 - ia_3 \\ a_2 + ia_3 & a_0 - a_1 \end{bmatrix}$
$-i\boldsymbol{\sigma}_2$	$\Phi \rightarrow \det(\Phi)\Phi^{-1}$	(4.3.14c)	$\begin{bmatrix} a_0 + ia_1 & a_2 - ia_3 \\ a_2 + ia_3 & a_0 - ia_1 \end{bmatrix}$

<sup>a</sup>Terms  $a_i$  ( $i = 0, 1, 2, 3$ ) are real numbers.

with  $ab + cd = 0$  and  $b^2 + d^2 = a^2 + c^2 = 1$ . Examples of unitary matrices that satisfy these constraints are  $\sigma_0$ ,  $\sigma_1$ ,  $\sigma_2$ , but not  $\sigma_3$ . The fact that  $\sigma_3$  violates these requirements may be interpreted as a signature of the breaking of the time-reversal symmetry. Another example of application lies in the detection of broken time-reversal symmetry in optical materials [71].

#### 4.3.2.2. Reciprocal Jones Medium Constraint

This subsection is devoted to the analysis of reciprocity and its physical consequence. The study of reciprocity requires a comparison of waves propagating in the forward and backward directions.

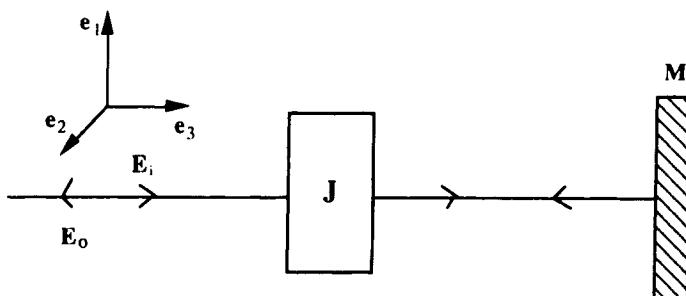
To examine the effect of reciprocity, we assume that magnetic fields are absent. We consider a Jones scattering matrix  $J(\mathbf{n}, \mathbf{n}')$  that describes how the field components of a plane wave, incident on an optical system, such as scattering medium, in the direction  $\mathbf{n}$ , are transformed in the far-field zone in the direction  $\mathbf{n}'$ . The fundamental reciprocity condition may be expressed in the form [71, 72]

$$J(\mathbf{n}, \mathbf{n}') = J^T(-\mathbf{n}', \mathbf{n}) \quad (4.3.15)$$

Note that for the reverse direction of the wave propagation, the Jones matrix of a rotator is  $R^T(\theta) = R(-\theta)$ , while for a compensator we have  $C^T(\delta) = C(\delta)$ .

For illustrative purposes, we consider the experiment displayed in Fig. 4.3.4. We use a right-handed Cartesian coordinate system that defines the laboratory coordinate system. The directions of light propagation for the forward and backward waves are specified by the unit vectors  $\mathbf{e}_3$  and  $-\mathbf{e}_3$ . It is convenient to use the same coordinate system for both directions of propagation, as shown in Fig. 4.3.4.

Let  $\mathbf{E}_i$  denote a plane wave of wave vector along  $\mathbf{e}_3$  incident on an optical medium. We assume that the medium is of the kind considered in Section 4.1



**Figure 4.3.4.** Notations related to the double passage of light through an optical system described by a Jones matrix, where  $\mathbf{e}_3$  is the unit vector in the direction of propagation of the incident plane wavefield and  $M$  is a perfect lossless isotropic mirror.

and can be characterized by the Jones matrix  $\mathbf{J}$ . After interacting with the optical system, the plane wave is retroreflected (i.e. retraces its path) at normal incidence by a perfect plane mirror and analyzed in the far-field zone. In general, it emerges from the system with its polarization modified. The Jones matrix is defined for a given direction of propagation of light; we denote by  $\mathbf{J}(\mathbf{e}_3)$  the Jones matrix for the direct propagation and  $\mathbf{J}(-\mathbf{e}_3)$  the Jones matrix when the direction of propagation is reversed. From Eq. (4.1.1), it follows that

$$\mathbf{E}_0(\mathbf{e}_3) = \mathbf{J}(\mathbf{e}_3)\mathbf{E}_i(\mathbf{e}_3) \quad (4.3.16)$$

For the reversed direction of propagation of light, we have

$$\mathbf{E}_0(-\mathbf{e}_3) = \mathbf{J}(-\mathbf{e}_3)\mathbf{E}_i(\mathbf{e}_3) = \mathbf{J}^T(\mathbf{e}_3)\mathbf{E}_i(-\mathbf{e}_3) \quad (4.3.17)$$

It has been noted previously that the transformation from a representation of the real rectangular basis in the incident coordinate system to the reflected system implies a Jones matrix of the form  $\mathbf{M} \equiv \boldsymbol{\sigma}_1$ , Eq. (4.1.57). Thus the output Jones vector may be written as

$$\mathbf{E}_0(\mathbf{e}_3) = \tilde{\mathbf{J}}\mathbf{E}_i(\mathbf{e}_3) \quad (4.3.18)$$

where

$$\tilde{\mathbf{J}} = \mathbf{J}^T \boldsymbol{\sigma}_1 \mathbf{J} \quad (4.3.19)$$

is the expression for the round-trip Jones matrix. Equivalently one may write

$$\tilde{\mathbf{J}} = \begin{bmatrix} J_{11}^2 - J_{21}^2 & J_{11}J_{12} - J_{21}J_{22} \\ J_{11}J_{12} - J_{21}J_{22} & J_{12}^2 - J_{22}^2 \end{bmatrix} \quad (4.3.20)$$

Having determined  $\tilde{\mathbf{J}}$ , we can ask for the condition of light extinction, that is,  $g = 0$ , after a round trip through the Jones medium [73]. We require that the following conditions be satisfied

$$J_{11}^2 - J_{21}^2 = J_{11}J_{12} - J_{21}J_{22} = J_{12}^2 - J_{22}^2 = 0 \quad (4.3.21)$$

leading to

$$|J_{11}| = |J_{21}|, \quad |J_{12}| = |J_{22}| \quad (4.3.22)$$

Before leaving this section, two comments are in order. First, the reciprocity relation Eq. (4.3.15) implies that a cascade of reciprocal Jones media is again a reciprocal Jones medium. It is a simple matter to prove that common optical devices such as rotators, polarizers, and compensators are reciprocal

elements. Second, we should note that Vansteenkiste and colleagues have obtained a number of reciprocity theorems that relate the polarization ellipticity at the output of an optical system to the polarization of the retroreflected light at the input [74]. These results have important practical applications for the use of polarization in remote sensing. A particularly noteworthy application has been for polarization control in a place to which experimentalists have no access, such as a vacuum chamber [74]. The reader is referred to Martinelli [75] and Bhandari [76] for discussions of other applications.

#### 4.3.2.3. Reciprocal Mueller Medium Constraint

We restricted our remarks up to this point to the case of optical systems characterized by Jones matrices. In view of the importance of the Mueller formalism to describe interactions of light with optical systems, we will now outline the relevant analysis to deal with reciprocity. Schönhofer and Kuball [77] have shown that the Mueller matrix for the reverse direction of light may be written in the form

$$\hat{\mathbf{M}} = \mathbf{OM}^T \mathbf{O}^{-1} \quad (4.3.23)$$

where  $\mathbf{M}$  is the Mueller matrix of the optical system and with

$$\mathbf{O} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4.3.24)$$

This expression is of importance in some investigations concerning randomly oriented particles, for example,  $\mathbf{M}_2$ , Eq. (4.3.5), satisfies the condition of reciprocity, Eq. (4.3.23).

## SECTION 4.4

---

# Random Media

Thus far our discussion has centered on optical media that are deterministic. We now extend our analysis to random media; specifically, their transformation matrices are described by an ensemble of realizations. The wave propagation in three-dimensional random media is accompanied by a number of particularly interesting features; among the dramatic effects observed have been the depolarization and enhanced backscattering of light from diffusively scattering systems as evinced by Refs. 78 through 82. The problem of wave propagation and scattering in random media has a long history in many areas of research, including optical propagation in oceans to investigate suspended sediments (oceanic hydrosols), and scattering of light by rough surfaces, to mention only a few of these phenomena.

Interest in this topic has been also stimulated by the considerable effort expanded on the study of speckle pattern correlations to retrieve useful information about the properties of a scattering medium. Speckle patterns arise when the local electric field is given by a sum of many contributions whose amplitudes and phases are randomly distributed. When a beam of light with uniform intensity migrates through a scattering medium, it undergoes a large number of scattering events from the inhomogeneities. The scattered waves interfere with each other, and, as a result, a highly irregular intensity pattern, with large intensity changes over short distances, is formed. The emerging intensity pattern is not uniform but is composed of many bright and dark spots called "speckles." If the specific arrangement of the inhomogeneities is static, only one microscopic realization of the disorder is being analyzed. If one measures only one realization, huge fluctuations can be observed. Thus the "speckle pattern" phenomenon refers to an intensity pattern formed, in the bulk of a disordered medium when a wave propagates through it, or by light reflected from a rough surface; it is a kind of "fingerprint" of a given configuration of scatterers in the medium. Note that the contributions leading to a speckle pattern change from one position to another in the plane of observation. If the scattering particles in the medium are moving, the speckle pattern will fluctuate in time. What I mean to suggest in these preliminary remarks is that the possible variety of behavior that occurs in the study of the interaction of radiation with a random medium is much richer than that for a strictly deterministic medium.

The aim of this section is to analyze how temporal or spatial fluctuations in a non-image-forming or scattering medium influence the output degree of polarization, the gain, and the output Stokes parameters of an incident quasimonochromatic wave of specified degree of polarization. For that purpose we exploit the statistical properties of light to probe the randomness in these complex systems.

The material treated in this chapter may be divided into two major topics. The first part discusses the effect of transmission of a plane wave by a temporally random medium. The second concerns the multiple elastic scattering of plane waves by isotropic ensembles of spherical dielectric particles. This refers to the situation in which each independent and randomly positioned scatterer (treated as a spherical particle) contributes to the radiant energy incident on the other scatterers. Here we present a detailed analysis of transmission and backscattering enhancement phenomena. To analyze the transport properties of plane waves in particulate media, we will consider the second-order statistics of the field. In the case of particles small compared to the wavelength of the incident wave, the dependence of the Stokes parameters of the emergent light as a function of the number of scattering events can be exactly evaluated. For Mie scattering, it is necessary to use numerical methods, specifically, Monte Carlo modeling, to calculate the Stokes parameters of the field. We discuss the calculation of properties such as the transport mean free path and the depolarization length. Comparison with results available from computer simulations demonstrates that the Monte Carlo model captures the trends in the transport properties of waves and yields realistic predictions for the depolarization length. To close this chapter, the enhanced backscattering of light from a medium with randomly positioned isotropic scatterers, which manifests as a narrow peak in the angular distribution of the intensity of the scattered light at scattering angles near  $180^\circ$ , is briefly discussed.

#### **4.4.1. JONES AND MUELLER MATRICES FOR TEMPORALLY RANDOM MEDIA**

The problem we will now consider is the characterization of the statistical behavior of a set of waves as they pass through a linear random medium. This question is critical for understanding systems with spatiotemporal disorder since field fluctuations carry information on the dynamics of these systems. The subject may be introduced in a variety of ways. This section is based on published work by Brosseau and Barakat [83]. For another approach, see, for example, Kim and coworkers [84]. Here it will be convenient to proceed in a similar manner as in Section 4.1 for deterministic media. Before turning to mathematical details, we will find it useful to retain a number of restrictions: (1) the far-field assumption permits us to describe the field as consisting of infinitely extended plane waves, (2) optical media are so optically thin that multiple reflections are negligible, (3) we consider only the influence of

temporal fluctuations in the medium and assume that the spatial fluctuations vary more slowly than the temporal fluctuations in the medium to be neglected in order to employ the usual Jones and Mueller matrix representation, and (4) we simplify our task by considering only forward scattering. We further note that the restriction to a plane parallel medium is not a stringent assumption since the radiance need be uniform over a surface for only a few optical mean free paths.

We have seen in Section 4.1 that the effect of a nonimage optical medium on the light is to relate the input and output coherency matrices by a congruent transformation, respectively  $\Phi_i$  and  $\Phi_o$ . Thus the problem of characterizing the statistics of scattered waves reduces to the determination of  $\Phi_o$ . Here we generalize that derivation to the case in which the optical medium is assumed to be temporally random. In regard to the incident plane wavefield, we will assume that it is Gaussian distributed. The wave is incident onto a slab-shaped nonmagnetic medium characterized by a Jones matrix that we decompose as a sum of a deterministic part  $J_D$  and a stationary ergodic zero-mean Gaussian (noise) process  $J_R$  so that

$$\langle J_R(t) \rangle = 0 \quad (4.4.1)$$

Consequently we have

$$\langle J(t) \rangle = J_D(t) \quad (4.4.2)$$

The reasons for our choice of a simple signal-plus-noise model for analyzing this problem are twofold: (1) this is a difficult problem, and solutions can be found only by limiting the degree of complexity of the analysis and (2) the response of a linear filter to noise is classic in circuit theory [14].

Then the input and output are related by an expression of the form

$$E_o(t; T) = J_D E_i(t; T) + J_R E_i(t; T) \quad (4.4.3)$$

If now we substitute from Eq. (4.4.3) into Eq. (4.1.7), we arrive at the relation

$$\begin{aligned} \langle E_o \times E_o^+ \rangle &= J_D \langle E_i \times E_i^+ \rangle J_D^+ + \langle J_R E_i \times E_i^+ J_R^+ \rangle + J_D \langle E_i \times E_i^+ J_R^+ \rangle \\ &\quad + \langle J_R E_i \times E_i^+ \rangle J_D^+ \end{aligned} \quad (4.4.4)$$

The last two terms vanish because they are linear in  $J_R$ , leaving only the second term with which to contend as the first term has already been studied in Section 4.1.1. The second term is evaluated by writing out the matrix and vector products, averaging, dividing by  $2T$ , and then passing to the limit  $T \rightarrow \infty$ . We now make the additional assumption that the fluctuations in the medium are much slower than the fluctuations in the incident field, so that these two kinds

of fluctuation are statistically independent. Consequently, the second term of the right-hand side of Eq. (4.4.4) may be expressed as

$$\Lambda = \lim_{T \rightarrow \infty} \frac{1}{2T} \langle \mathbf{J}_R \mathbf{E}_i \times \mathbf{E}_i^+ \mathbf{J}_R^+ \rangle \quad (4.4.5)$$

whose respective matrix elements  $\Lambda_{ij}$  are

$$\begin{aligned} \Lambda_{11} &= \langle |\tau_{11}|^2 \rangle \Phi_{11} + \langle \tau_{11} \tau_{12}^* \rangle \Phi_{12} + \langle \tau_{12} \tau_{11}^* \rangle \Phi_{21} + \langle |\tau_{12}|^2 \rangle \Phi_{22} \\ \Lambda_{12} &= \langle \tau_{11} \tau_{21}^* \rangle \Phi_{11} + \langle \tau_{11} \tau_{22}^* \rangle \Phi_{12} + \langle \tau_{21} \tau_{12}^* \rangle \Phi_{21} + \langle \tau_{12} \tau_{22}^* \rangle \Phi_{22} \\ \Lambda_{21} &= \langle \tau_{21} \tau_{11}^* \rangle \Phi_{11} + \langle \tau_{21} \tau_{12}^* \rangle \Phi_{12} + \langle \tau_{22} \tau_{11}^* \rangle \Phi_{21} + \langle \tau_{22} \tau_{12}^* \rangle \Phi_{22} \\ \Lambda_{22} &= \langle |\tau_{21}|^2 \rangle \Phi_{11} + \langle \tau_{21} \tau_{22}^* \rangle \Phi_{12} + \langle \tau_{22} \tau_{21}^* \rangle \Phi_{21} + \langle |\tau_{22}|^2 \rangle \Phi_{22} \end{aligned} \quad (4.4.6)$$

where we have denoted the matrix elements of  $\mathbf{J}_R$  by  $\tau_{jk}$ . The output coherency matrix reduces in the far-field region to

$$\Phi_o = \mathbf{B} + \Lambda \quad (4.4.7)$$

where we have set

$$\mathbf{B} = \mathbf{J}_D \Phi_i \mathbf{J}_D^+ \quad (4.4.8)$$

The expression (3.1.126) for the average Stokes parameters is still valid. On carrying out the necessary computations, we have

$$\langle S_i \rangle = \sum_{k=0}^3 D_k^i \langle S_k \rangle \quad (4.4.9)$$

A list of the  $D_k^i$  coefficients is given in Appendix J. The fluctuations of the amplitudes and phases that arise in a wave traversing a randomly irregular medium contain information both about the source of the wavefield and the random medium itself. Note that if the matrix elements of  $\mathbf{J}_R$  are themselves uncorrelated, then only  $\langle S_0 \rangle$  and  $\langle S_1 \rangle$  are influenced by temporal fluctuations.

As before, we can ask for the gain of the medium and the output degree of polarization

$$g = \frac{1}{2} \sum_{k=0}^3 D_k^0 \langle \sigma_k \rangle_i \quad (4.4.10)$$

where, as before,  $\langle \sigma_k \rangle_i$  denote the input normalized Stokes parameters. A related result is that the measurement of the gain for various input states of polarization yields information on the second-order moments of the  $\tau$  values.

As in the deterministic case, the maximum and minimum gain are respectively given by

$$g_{\max} = D_0^0 + P_i \sum_{k=0}^3 D_k^0 \langle S_k \rangle_i \quad (4.4.11a)$$

and

$$g_{\min} = D_0^0 - P_i \sum_{k=0}^3 D_k^0 \langle S_k \rangle_i \quad (4.4.11b)$$

It is instructive to compare the formulas (4.4.11a, b) with the deterministic case, Eqs. (4.1.18a, b). Of particular interest is the case where the  $\tau$  values are uncorrelated; we then have, from Eqs. (J.16)

$$\begin{aligned} D_0^0 &= J_0 + K_0 & D_1^0 &= J_1 + K_1 \\ D_2^0 &= J_2 & D_3^0 &= J_3 \end{aligned} \quad (4.4.12)$$

where  $K_0$  and  $K_1$  represent the contributions to  $D_0^0$  and  $D_1^0$  arising from the fluctuations in the medium. If further the fluctuations are small, then  $J_0 \gg K_0$ ,  $J_1 \gg K_1$  and

$$g_{\max(\text{random})} \cong g_{\max(\text{deterministic})} + \frac{1}{2} K_0 + \frac{1}{2} K_1 J_1 P \quad (4.4.13)$$

Although  $K_0 \geq 0$ , both  $J_1$  and  $K_1$  can be negative, allowing for the possibility that  $g_{\max(\text{random})} < g_{\max(\text{deterministic})}$ .

We next derive an expression of the input degree of polarization. For this purpose we take the determinant of both sides of Eq. (4.4.7). The right-hand side can be decomposed as

$$\det(\mathbf{B} + \boldsymbol{\Lambda}) = \det(\mathbf{B}) + \det(\boldsymbol{\Lambda}) + (B_{11}\Lambda_{22} + B_{22}\Lambda_{11}) - (B_{12}\Lambda_{12}^* - B_{12}^*\Lambda_{12}) \quad (4.4.14)$$

Substituting Eq. (4.4.8) into Eq. (4.4.14) we have

$$\begin{aligned} \langle S_0 \rangle_o^2 (1 - P_o^2) &= \langle S_0 \rangle_i^2 (1 - P_i^2) |\det(\mathbf{J}_D)|^2 \\ &\quad + 4 \{ \det(\boldsymbol{\Lambda}) + (B_{11}\Lambda_{22} + B_{22}\Lambda_{11}) - (B_{12}\Lambda_{12}^* - B_{12}^*\Lambda_{12}) \} \end{aligned} \quad (4.4.15)$$

Hence we see from Eq. (4.4.15) how the output degree of polarization depends on the statistical characteristic of the random medium and of the polarization state of the incident light. The main feature emerging from this formula is that it is no longer true that polarized light passing through the optical system remains totally polarized as was required in the deterministic case.

The situation discussed above in the context of the Jones matrix has an equivalent description by means of random Mueller-Jones matrices. First consider the sample realization connecting the output Stokes vector  $\mathbf{S}_o$  and the input Stokes vector  $\mathbf{S}_i$

$$\mathbf{S}_o = \mathbf{M}_j \mathbf{S}_i, \quad (4.4.16)$$

where  $\mathbf{M}_j$  itself is random. Since the fluctuations of  $\mathbf{M}_j$  and of  $\mathbf{S}_i$  are uncorrelated, we arrive at the relation

$$\langle \mathbf{S}_o \rangle = \langle \mathbf{M} \mathbf{S}_i \rangle = \langle \mathbf{M} \rangle \langle \mathbf{S}_i \rangle \quad (4.4.17)$$

The explicit representation of random  $\mathbf{M}_j$  in terms of random  $\mathbf{J}$  follows directly from Eq. (4.1.139). Thus a sample realization of  $\mathbf{M}_j$  is

$$\mathbf{M}_j = \mathbf{A}(\mathbf{J}_D \otimes \mathbf{J}_D^*)\mathbf{A}^{-1} + \mathbf{A}(\mathbf{J}_R \otimes \mathbf{J}_R^*)\mathbf{A}^{-1} + \mathbf{A}(\mathbf{J}_D \otimes \mathbf{J}_R^*)\mathbf{A}^{-1} + \mathbf{A}(\mathbf{J}_R \otimes \mathbf{J}_D^*)\mathbf{A}^{-1} \quad (4.4.18)$$

The last two terms, which are linear in  $\mathbf{J}_R$ , vanish on averaging, leaving

$$\langle \mathbf{M}_j \rangle = \mathbf{A}(\mathbf{J}_D \otimes \mathbf{J}_D^*)\mathbf{A}^{-1} + \mathbf{A}\langle \mathbf{J}_R \otimes \mathbf{J}_R^* \rangle \mathbf{A}^{-1} \equiv \mathbf{M}_j^{(D)} + \mathbf{M}_j^{(R)} \quad (4.4.19)$$

Consequently there are two independent contributions to  $\langle \mathbf{M}_j \rangle$ , one from the deterministic term and the other from the fluctuations in the medium. The gain reads

$$g = \sum_{k=0}^3 \mu_{0k} \langle \sigma_k \rangle_i \quad (4.4.20)$$

where  $\mu_{0k}$  are the matrix elements of the first row of  $\langle \mathbf{M}_j \rangle$ . These matrix elements, in turn, are equal to the sum of the matrix elements of the first row of  $\mathbf{M}_j^{(D)}$  and  $\mathbf{M}_j^{(R)}$ , which are simply the  $D_k^0$  in Eq. (4.4.9). Thus the gain of the random medium as described by either the random Jones matrix or the random Mueller-Jones matrix is the same, as expected. As before, one can show that the output degree of polarization  $P_o$  in the  $\langle \mathbf{M}_j \rangle$  description is again given by Eq. (4.4.15).

We now outline a generalization of trace equation (4.1.49) for a random Mueller-Jones matrix. Let us begin with Eq. (4.4.18), which can also be written as

$$\mathbf{M}_j = \Theta_D + \Theta_R + \Theta_3 + \Theta_4 \quad (4.4.21)$$

where

$$\Theta_D = A(J_D \otimes J_D^*)A^{-1} \quad (4.4.22a)$$

$$\Theta_R = A(J_R \otimes J_R^*)A^{-1} \quad (4.4.22b)$$

$$\Theta_3 = A(J_D \otimes J_R^*)A^{-1} \quad (4.4.22c)$$

$$\Theta_4 = A(J_R \otimes J_D^*)A^{-1} \quad (4.4.22d)$$

We are interested in writing an expression of  $\langle M_J^T M_J \rangle$ . Taking the product of  $M_J^T (\equiv M_J^+)$  since is real-valued) with  $M_J$ , from Eq. (4.4.21), averaging, and after a rearranging of the terms, results in

$$\begin{aligned} \langle M_J^T M_J \rangle &= \Theta_D^+ \Theta_D + \langle \Theta_R^+ \Theta_R \rangle + \langle \Theta_3^+ \Theta_3 \rangle + \langle \Theta_4^+ \Theta_4 \rangle + \langle \Theta_D^+ \Theta_R + \Theta_R^+ \Theta_D \rangle \\ &\quad + \langle \Theta_D^+ \Theta_3 + \Theta_3^+ \Theta_D \rangle + \langle \Theta_D^+ \Theta_4 + \Theta_4^+ \Theta_D \rangle + \langle \Theta_R^+ \Theta_3 + \Theta_3^+ \Theta_R \rangle \\ &\quad + \langle \Theta_R^+ \Theta_4 + \Theta_4^+ \Theta_R \rangle + \langle \Theta_3^+ \Theta_4 + \Theta_4^+ \Theta_3 \rangle \end{aligned} \quad (4.4.23)$$

From the Gaussian property of the fluctuations and the fact that they are zero-mean, it follows that the sixth to the ninth terms of the right-hand side of Eq. (4.4.23) are zero. The problem thus reduces to a determination of the nonzero terms. From the usual properties of the Kronecker product, we get

$$\langle \Theta_R^+ \Theta_R \rangle = A(J_R^+ J_R \otimes (J_R^+ J_R)^*)A^{-1} = A(\Omega_R \otimes (\Omega_R)^*)A^{-1} \quad (4.4.24a)$$

$$\langle \Theta_3^+ \Theta_3 \rangle = A(\Omega_D \otimes (\Omega_R)^*)A^{-1} \quad (4.4.24b)$$

$$\langle \Theta_4^+ \Theta_4 \rangle = A(\Omega_R \otimes (\Omega_D)^*)A^{-1} \quad (4.4.24c)$$

$$\langle \Theta_D^+ \Theta_R + \Theta_R^+ \Theta_D \rangle = A(\Omega_{DR} \otimes (\Omega_{DR})^* + \Omega_{DR}^+ \otimes (\Omega_{DR}^+)^*)A^{-1} \quad (4.4.24d)$$

$$\langle \Theta_3^+ \Theta_4 + \Theta_4^+ \Theta_3 \rangle = A(\Omega_{DR} \otimes (\Omega_{DR}^+)^* + \Omega_{DR}^+ \otimes \Omega_{DR}^*)A^{-1} \quad (4.4.24e)$$

where we have set

$$\Omega_D \equiv J_D^+ J_D \quad (4.4.25a)$$

$$\Omega_R \equiv J_R^+ J_R \quad (4.4.25b)$$

$$\Omega_{DR} \equiv J_D^+ J_R = \Omega_{RD}^+ \quad (4.4.25c)$$

Then taking the trace of Eq. (4.4.23) and remembering that the trace is invariant under cyclic permutation of factors, we can rewrite the trace in the form

$$\begin{aligned} \text{tr} \langle M_J^T M_J \rangle &= \text{tr}(\Omega_D \otimes \Omega_D^*) + \text{tr}(\Omega_R \otimes \Omega_R^*) + \text{tr}(\Omega_D \otimes \Omega_R^*) + \text{tr}(\Omega_R \otimes \Omega_D^*) \\ &\quad + \text{tr}(\Omega_{DR} \otimes (\Omega_{DR})^* + \Omega_{DR}^+ \otimes (\Omega_{DR}^+)^*) + \text{tr}(\Omega_{DR} \otimes (\Omega_{DR}^+)^* + \Omega_{DR}^+ \otimes \Omega_{DR}^*) \end{aligned} \quad (4.4.26)$$

Using the fact that the trace and expectation operators commute, we also note that  $\text{tr}\langle \mathbf{M}_J^T \mathbf{M}_J \rangle = \langle \text{tr}(\mathbf{M}_J^T \mathbf{M}_J) \rangle$ . The end result is obtained with Kronecker algebra:

$$\begin{aligned} \text{tr}\langle \mathbf{M}_J^T \mathbf{M}_J \rangle &= 4\langle \mu_{00}^2 \rangle = (\text{tr}(\mathbf{J}_D^+ \mathbf{J}_D))^2 + (\text{tr}(\mathbf{J}_R^+ \mathbf{J}_R))^2 \\ &\quad + 4\langle (\text{Re}\{\text{tr}(\mathbf{J}_D^+ \mathbf{J}_R)\})^2 \rangle + 2\text{tr}(\mathbf{J}_D^+ \mathbf{J}_D)\langle \text{tr}(\mathbf{J}_R^+ \mathbf{J}_R) \rangle \end{aligned} \quad (4.4.27)$$

The expression for  $\text{tr}(\langle \mathbf{M}_J \rangle^T \langle \mathbf{M}_J \rangle)$  can be evaluated in the same fashion as above and is written in final form as

$$\text{tr}(\langle \mathbf{M}_J \rangle^T \langle \mathbf{M}_J \rangle) = (\text{tr}(\mathbf{J}_D^+ \mathbf{J}_D))^2 + 2\langle |\text{tr}(\mathbf{J}_D^+ \mathbf{J}_R)|^2 \rangle + \text{tr}(\langle \mathbf{J}_R \otimes \mathbf{J}_R^+ \rangle^+ \langle \mathbf{J}_R \otimes \mathbf{J}_R^* \rangle) \quad (4.4.28)$$

Equations (4.4.27) generalizes Eq. (4.1.148) to a Mueller–Jones medium that is represented by an ensemble of realizations. It is sufficiently interesting to warrant some comments. When  $\mathbf{J}_R \equiv 0$  (deterministic medium), or under a wider class of conditions when the last three terms of the right-hand side of Eqs. (4.4.27) add up to zero, we recover Eq. (4.1.148):

$$\text{tr}(\mathbf{M}_J^T \mathbf{M}_J) = |\text{tr}(\Omega_D)|^2 = (\text{tr}(\mathbf{J}_D^+ \mathbf{J}_D))^2 \quad (4.4.29)$$

Since the (0-0) entry of  $\mathbf{M}_J$  now is a random process, the different averages in Eq. (4.4.27) characterize the second-order fluctuations of the gain for unpolarized light. One can obtain these averages experimentally by measuring the first-order moments of  $\mu_{00}$ , specifically,  $\langle \mu_{00} \rangle = \frac{1}{2}(\text{tr}(\mathbf{J}_D^+ \mathbf{J}_D) + \langle (\text{tr}(\mathbf{J}_R^+ \mathbf{J}_R)) \rangle + \langle \text{Re}\{\text{tr}(\mathbf{J}_D^+ \mathbf{J}_R)\} \rangle)$  and  $\langle \mu_{00}^2 \rangle$  given by Eq. (4.4.27).

Since the second and third terms of the right-hand side of Eq. (4.4.28) are always nonnegative, we obtain the following inequality:

$$\text{tr}(\langle \mathbf{M}_J \rangle^T \langle \mathbf{M}_J \rangle) \geq (\text{tr}(\mathbf{J}_D^+ \mathbf{J}_D))^2 = \text{tr}((\mathbf{M}_J^{(D)})^T \mathbf{M}_J^{(D)}) \quad (4.4.30)$$

This expression is important because it shows that the trace condition Eq. (4.1.148), takes the form of an inequality in general. To this point we have imposed that the incident wavefield is Gaussian distributed. Were we not to make this assumption, we would lose the simplifications in Eq. (4.4.23), and the problem would become intractable. As our final comment, we point out that there seems to be very little that we can say, in general, about the random Mueller matrix.

#### 4.4.2. MULTIPLE SCATTERING BY A SPATIALLY RANDOM MEDIUM

Many common observations, for example, the variation of color of the sky or the darkening of sand on wetting, cannot be explained by single-scattering arguments. Multiple scattering of waves by a collection of particles has been of

interest in many areas of soft condensed matter (e.g., colloids or collections of macromolecules) as well as technology (e.g., paint coatings, emulsions), in an attempt to quantitatively study the collective structure, interparticle correlation, or morphology. Historically, theoretical solutions for the scattering characteristics of three-dimensional objects have involved approximations for special classes of scatterers. For example, the Rayleigh solution is applicable to objects whose size is small relative to a given wavelength. This approximation permits us to treat the small particle as an electric dipole. The incident oscillating electric field is essentially constant over the entire particle. In this case the particle becomes invisible because it is too small to scatter very much radiation. The solution to the problem of light scattering by objects whose size is of the order of the wavelength or larger, is known as the *Lorentz–Mie solution*.

Propagation and scattering of electromagnetic waves in an inhomogeneous medium depends critically on the ratio between the wavelength and the scale lengths of the inhomogeneities. We emphasize at the outset that wave transport through a medium with randomly positioned scatterers can be characterized by a set of significant scale lengths. The first scale is the thickness  $d$  of the optical medium. The second scale is the elastic mean free path  $l \equiv 1/\phi\sigma$ , the average length the wave travels before it suffers an elastic collision. Here  $\phi$  is the concentration of scatterers and  $\sigma$  is a scattering cross section. The third scale worth considering is the transport mean free path  $l^*$ , which is defined as the average distance over which momentum transfer becomes uncorrelated; thus the wave propagates a distance of the order of  $l^*$  before it completely forgets about its initial direction of propagation. The fourth scale is the wavelength  $\lambda$ . The fifth scale is the size of the scatterers, namely,  $a$ . Note that in the general case of a polydisperse system, the size, shape, and refractive-index distributions of the scatterers are characterized by a distribution of lengths. We will find it useful to introduce a dimensionless size parameter  $qa$ , where  $a$  is the radius of the particles and  $q$  is the wavenumber of the wave. Waves with short wavelength see a smoothly varying medium, while long waves essentially do not feel the inhomogeneities. We would like to outline here why the sixth scale—the length  $\xi$  of the path over which a polarized wave becomes depolarized—emerges from this analysis and how it depends on whether it is initially linearly or circularly polarized, on the size on the particles, and on the anisotropy of the diffusers that scatter light.

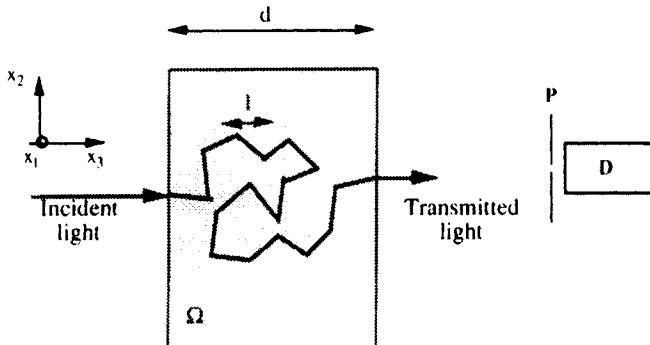
If  $d \leq l^*$ , the inhomogeneities in the medium give rise to only weak elastic scattering. Consequently most of the incident wave is unscattered on emerging from the sample. The single-scattering events dominate; thus the scattered wave can be conveniently described by the Born approximation in standard scattering theory. At the opposite multiple scattering regime  $d \gg l^*$  (but still in the weak-localized regime  $\lambda \ll l^*$ ), a wave scatters many times before emerging from the sample. The unscattered component of the transmitted wave is exponentially attenuated by a factor  $\exp(d/l^*)$ . The typical number of scattering events by a particular wave propagating across the sample is  $\sim(d/l^*)^2 \gg 1$  [80].

This section is presented in three parts. Not many problems of wave scattering are in any sense exactly soluble. But there does exist one completely soluble problem that has considerable physical meaning. This is Rayleigh scattering. In the first part we discuss general features of the analysis of the transmission of a beam of light that is normally incident on a particulate medium constituted of small randomly oriented spherical particles (Rayleigh limit) that gives closed-form expressions for the degree of polarization. In the second part of this section, a Monte Carlo modeling of the wave propagation in three-dimensional media composed of independently scattering particles is carried out over a range of particle dimensions that span values from  $qa \ll 1$  to  $qa \cong 15$ . We also include comparisons of our numerical simulations with experimentally obtained far-field scattering data obtained from polystyrene spheres. Then, in Section 4.4.2.3, we describe enhanced backscattering of polarized light from a medium with randomly positioned isotropic scatterers. The thrust of this section is on the principle underlying the physical ideas and not on applications of these concepts to practical useful applications.

#### 4.4.2.1. Rayleigh Scattering

In this section we want to study the propagation of an incident pure state of polarization in a medium with randomly positioned particles such that multiple-scattering effects cannot be neglected, as was implicitly done earlier in Section 4.3.1. We present a self-contained review of the physics behind the scattering of electromagnetic radiation from a pointlike suspension [85]. The derivation here focuses on the problem of determining the full Stokes vector for multiply scattered wave.

Let us suppose that a quasimonochromatic, of mean frequency  $\nu_0$ , plane wave is incident onto the left side of a scattering three-dimensional random medium that occupies a finite volume  $\Omega$  in free space, as displayed schematically in Fig. 4.4.1. The output wave intensity pattern will be a complicated speckle pattern. To describe the scattering of a polarized lightwave, we use a right-handed Cartesian coordinate system, referred to as the *laboratory reference frame*. This figure shows a possible path of a wave entering normally to the system. A useful physical picture for the propagation of the wave as it enters the sample is one in which a wave undergoes a random walk. Each trajectory is composed of straight-line segments and sudden interruptions that randomly change the wave's propagation direction. The average length of each random typical step is the mean free path  $l$ . In the weakly scattering regime, namely,  $\lambda \ll l^*$ , the wave intensity satisfies the classic diffusion equation. Then, for distances that are much larger than  $l^*$  beyond which the direction of light propagation is randomized, light transport can be regarded as a diffusion process with diffusion constant  $D = \frac{1}{3}vl^*$  where  $v$  is the transport velocity, or the speed of light in the medium. For pointlike scatterers,  $v$  is equal to the phase velocity, which is approximately equal to the velocity of light divided by the index of refraction. It is worth mentioning that, unlike the transport mean free path  $l^*$ , which is obtained experimentally from steady-state measurements,



**Figure 4.4.1.** Schematic diagram and notation relating to the wave propagation through a random scattering medium. A plane wave is normally incident on a nonabsorbing random medium, of volume  $\Omega$  and of thickness  $d \gg l$ , consisting of uncorrelated spherical pointlike particles (Rayleigh scattering). Typical scattering path executing zigzag random walk through the medium (propagating “channel”). The mean free path  $l$  is the typical step size. A single speckle of transmitted light is imaged with a pinhole  $P$  and is monitored with a photodetector  $D$ .

the diffusion constant  $D$  is obtained from dynamical measurements. The density of scatterers must be small enough to allow the weak scattering approximation to be valid.

Besides being nonabsorbing, the scattering medium is assumed to be time-invariant, nonmagnetic, and spatially nondispersive and such that the spatial fluctuations of its dielectric susceptibility  $u_{ij}(\mathbf{R})$  tensor are statistically homogeneous and stationary in space (at least in the wide sense). The incident and scattered beams are normal to the surfaces of the scattering medium, and the coordinates system lies parallel to the slab faces. Typical realization of such medium would be a collection of discrete pointlike, scattering centers whose size is very small compared to the wavelength of the scattered radiation (i.e.,  $qa \ll 1$ ). This approximation permits the small sphere to be treated as a dipolar oscillator with its polarizability determined by the optical constants of the particle.

We also assume that the temporal fluctuations of the scatterers are sufficiently slow relative to the period of the field oscillations that the scattering medium behaves as if it were essentially time-invariant (i.e., adiabatic approximation). The usual boundary conditions require continuity of the magnetic field  $\mathbf{H}$  and tangential electric field at every discontinuity surface. From the preceding assumptions, we may characterize the dielectric susceptibility of the three-dimensional medium by  $u_{ij}(\mathbf{R}) = u(\mathbf{R})\delta_{ij}$ , of zero mean and white-noise correlation function:

$$\langle u(\mathbf{R}_1)u(\mathbf{R}_2) \rangle = \begin{cases} u\delta(\mathbf{R}_1 - \mathbf{R}_2) & \text{when } \mathbf{R}_1 \in \Omega, \mathbf{R}_2 \in \Omega \\ 0 & \text{otherwise} \end{cases} \quad (4.4.31)$$

where  $u$  is a constant that is a measure of the scattering potential. Let us add a further condition. We will consider only weak disorder such that the elastic mean free path  $l \equiv 6\pi/uq_0^4$  is much larger than the wavelength of the radiation (i.e.  $q_0 l \gg 1$ ),  $q_0$  is the free-space wavenumber associated with the frequency  $\nu_0$ : consequently, the wavefield propagation may be described by a classic diffusion process. Finally, we do assume that the fluctuations of the medium and the fluctuations of the incident field are statistically independent. These restrictions present no very severe difficulties to experimental practice.

Now we are ready to begin. All information about an elastic scattering process is contained in the 16-element Mueller matrix. A number of restrictions are placed at the outset on the form of the  $\mathbf{M}$  matrix depending on the symmetry and reciprocity requirements. In view of the symmetry arguments presented in this Section 4, and the fact that multiple scattering is assumed to be free of absorption, the general form of the Mueller matrix  $\mathbf{M}$  can be written as

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & m_{11} & 0 & 0 \\ 0 & 0 & m_{22} & 0 \\ 0 & 0 & 0 & m_{33} \end{bmatrix} \quad (4.4.32)$$

The problem now shifts to the explicit calculation of the dependence of the degree of polarization on function of the number of scattering events when the medium fulfills the assumptions stated above. The Mueller matrix elements  $m_{ii}(n)$  can be calculated explicitly by the Bethe–Salpeter equation handled in the ladder approximation. Within this approach the problem of evaluating the coherency matrix reduces to a matrix eigenvalue problem. Interested readers may refer to Appendix J for full details of the derivation. However, it is worthy to note that the Bethe–Salpeter equation under the ladder approximation of uncorrelated discrete scatterers reduces to Chandrasekhar's radiative transfer equation [59]. We emphasize that this derivation is valid over distances greater than the mean free path.

Next we consider a pure state of polarization of unit intensity incident normally on the slab-shaped medium; its Stokes vector is expressed as follows:

$$\mathbf{S}_i = \begin{bmatrix} \langle S_0 \rangle = \langle |E_1|^2 \rangle + \langle |E_2|^2 \rangle = 1 \\ \langle S_1 \rangle = \langle |E_1|^2 \rangle - \langle |E_2|^2 \rangle \\ \langle S_2 \rangle = \langle E_1^* E_2 + E_1 E_2^* \rangle \\ \langle S_3 \rangle = i \langle E_1^* E_2 - E_1 E_2^* \rangle \end{bmatrix} \quad (4.4.33)$$

For a weak scattering limit, the linear response of the scattering medium is determined by the ensemble-averaged covariance satisfying the Bethe–Salpeter equation [Eq. (J.1)]. Following this analysis, we obtain the expression for the

output Stokes vector:

$$\mathbf{S}_o = \begin{bmatrix} \langle |G_{11}|^2 \rangle + \langle |G_{12}|^2 \rangle \\ \langle S_1 \rangle (\langle |G_{11}|^2 \rangle - \langle |G_{12}|^2 \rangle) \\ \langle S_2 \rangle (\langle G_{22}^* G_{11} \rangle + \langle G_{21}^* G_{12} \rangle) \\ \langle S_3 \rangle (\langle G_{22}^* G_{11} \rangle - \langle G_{21}^* G_{12} \rangle) \end{bmatrix} \quad (4.4.34)$$

where the  $G_{ij}$  terms denote the retarded Green's function [Eq. (J.13)] and the suffices 1, 2 label components with respect to the Cartesian coordinate system chosen. Since absorption is ignored, we normalize  $\mathbf{S}_o$  with respect to  $\langle |G_{11}|^2 \rangle + \langle |G_{12}|^2 \rangle$ . If we make use of the relations (J.16a-d) we see at once that  $\mathbf{S}_o$  may be expressed as

$$\mathbf{S}_o = \begin{bmatrix} 1 \\ \langle S_1 \rangle m_{11}(n) \\ \langle S_2 \rangle m_{11}(n) \\ \langle S_3 \rangle m_{33}(n) \end{bmatrix} \quad (4.4.35)$$

with  $m_{11}(n) = m_{22}(n) = 3(\frac{7}{10})^n / [2 + (\frac{7}{10})^n]$  and  $m_{33}(n) = 3(\frac{1}{2})^n / [2 + (\frac{7}{10})^n]$ , where  $n + 1$  is the number of scattering events. It is readily verified from Eq. (4.4.35) that the Mueller matrix of the scattering medium has the kind of symmetry we expect from (4.4.32). Having found the form of the Stokes vector  $\mathbf{S}_o$ , we might naturally wonder what form the output degree of polarization has. Under these circumstances, a straightforward calculation of the degree of polarization  $P_o$  of the emerging wave results in the expression

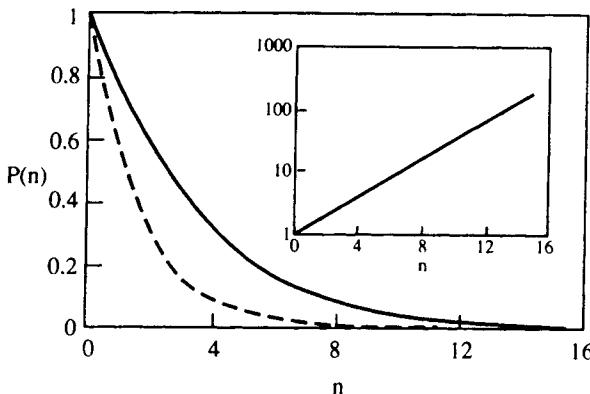
$$P_o = m_{11}(n) \{ \langle S_1 \rangle^2 + \langle S_2 \rangle^2 + \langle S_3 \rangle^2 (\frac{5}{7})^{2n} \}^{1/2} \quad (4.4.36)$$

which involves three independent parameters.

As it stands, Eq. (4.4.36) shows that the output degree of polarization will, in general, differ from the input degree of polarization because of the effect of the scattering medium. These equations are in accordance with the fact that single scattering (i.e.,  $n = 0$ ) by pointlike particles having spherical symmetry preserves the state and degree of polarization. Now it is instructive to consider several special cases. For instance, an input linear polarization state has for output Stokes vector

$$\begin{bmatrix} 1 \\ m_{11}(n) \\ 0 \\ 0 \end{bmatrix}$$

and a degree of polarization  $P_o = m_{11}(n)$  that is a monotonically decreasing function of the number of scattering. Similarly, for an input right-handed



**Figure 4.4.2.** Degree of polarization of scattered light as a function of the number of scatterings  $n$  [i.e., Eq. (4.4.36)] for an input pure state of linear parallel polarization (solid line), right circular polarization (dashed line). The inset depicts the dependence of the normalized parameter  $v(n)/v(0)$  on  $n$ .

circular polarization state, one gets

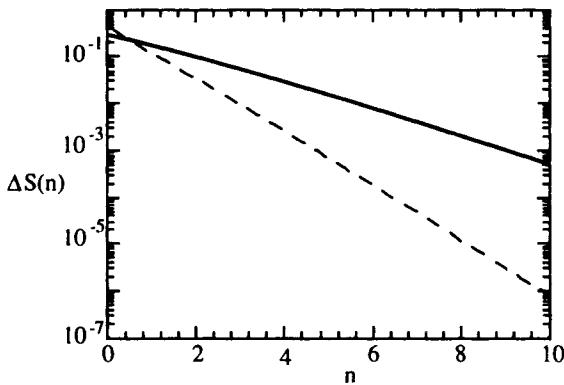
$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ m_{33}(n) \end{bmatrix}$$

and  $P_o = m_{33}(n)$ . Then, the two Mueller matrix elements  $m_{11}$  and  $m_{33}$  have clear physical meanings. Curves showing the behavior of the degree of polarization of the output wave are shown in Fig. 4.4.2.

The parameter  $v = [(\langle S_1 \rangle^2 + \langle S_2 \rangle^2)^{1/2}] / \langle S_3 \rangle$ , which represents the ratio of linear to circular polarization has been also displayed in Fig. 4.4.2. The exponential increase indicates that as the degree of polarization decreases, the polarization ellipse flattens toward the major axis. This fact is also expressed by saying that depolarization of linearly polarized light requires more scattering events than are required for a circularly polarized light (by a factor  $\approx 2$ ).

To further discuss the physical significance of these results, we have also plotted the variation of the entropy production  $\Delta S(n) = S(P(n+1)) - S(P(n))$  against the number of scattering events (Fig. 4.4.3). As can be seen,  $\Delta S(n)$  is well represented by an exponential decay  $\Delta S(n) = a \exp(-bn)$ , with  $a$  and  $b$  depend on the particular state of polarization. For large values of  $n$  (say,  $n \geq 10$ ), the entropy of radiation is unaffected by further scattering; it defines the steady state of maximum entropy [ $S(P=0) = \ln 2$ ] attainable by multiple scattering.

We observe that Rayleigh scattering depolarizes any arbitrary incident Stokes vector: in the limit of large numbers of scattering events, Eq. (4.4.32)

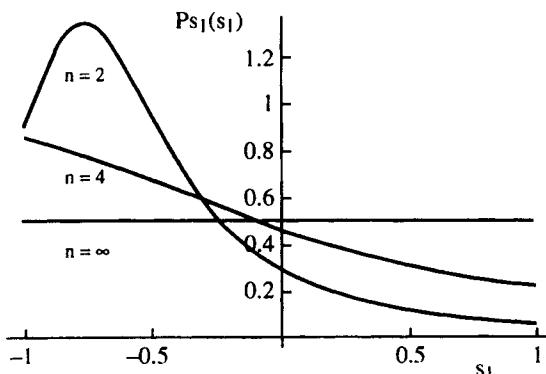


**Figure 4.4.3.** Entropy production  $\Delta S(n) = S(P(n+1)) - S(P(n))$  as a function of the number of scattering events. Same symbols as in Fig. 4.4.2.

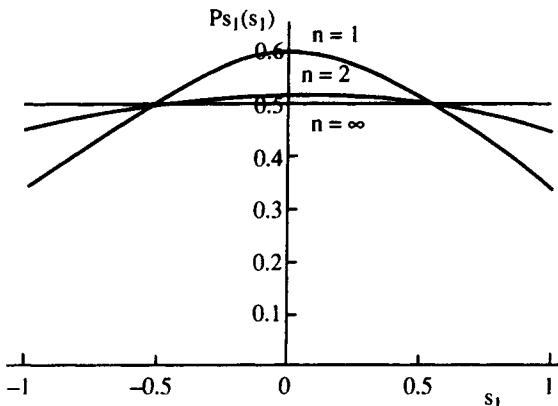
approaches the Mueller matrix of an ideal depolarizer, [Eq. (4.1.47)]:

$$\lim_{n \rightarrow \infty} \mathbf{M}_n = \mathbf{D} \quad (4.4.37)$$

An extension of the analysis to include the statistics of the normalized Stokes parameters is of some interest. Because of the large number of statistically independent contributions, the scattered field is Gaussian-distributed by virtue of the central-limit theorem. Here we are interested in correlating the asymmetry polarization effect to the behavior of the PDF of  $s_1$  [86]. Numerical values of  $p_{s_1}(s_1)$  are shown, for different values of  $n$ , in Fig. 4.4.4 for



**Figure 4.4.4.** PDF of  $s_1$  as function of  $s_1$  for different values of the number of scattering events  $n$ . This situation corresponds to light transmission through a multiple-scattering medium composed of uncorrelated, optically inactive spherical pointlike particles and a case of incident linearly polarized light ( $\langle S_1 \rangle / \langle S_0 \rangle = -0.5$ ).



**Figure 4.4.5.** PDF of  $s_1$  as function of  $s_1$  for different values of the number of scattering events  $n$ . This situation corresponds to light transmission through a multiple-scattering medium composed of uncorrelated, optically inactive spherical pointlike particles and a case of incident circularly polarized light ( $\langle S_1 \rangle = 0$ ).

an incident linearly polarized state ( $\langle \sigma_1 \rangle_i = -0.5$ ) and in Fig. 4.4.5 for an incident circularly polarized state ( $\langle \sigma_1 \rangle_i = 0$ ). Note how the PDF becomes flat as  $n$  is increased; ultimately  $\lim_{n \rightarrow \infty} p_{s_1}(s_1) = \frac{1}{2}$ . This is also in conformity with the preceding analysis of depolarization by multiple scattering in terms of nonnormalized Stokes parameters.

An important question is whether and when the assumption that light is a Gaussian random process is justified. It is certainly justified when  $n$  is large. However, if there is only a small number of scattering events, the scattered light is no longer Gaussian because the central limit cannot be invoked. If the light is non-Gaussian, the second-order description of the field is no longer sufficient to characterize the fluctuations [85].

In closing this subsection, it is also interesting to note that a related approach—the technique of diffusing-wave spectroscopy (DWS)—can be used to obtain information about the dynamics of the particles scattering the wave from the study of the temporal fluctuations of scattered light intensity. In a DWS experiment, one determines the (normalized) time correlation function of the scattered intensity in order to characterize density correlations and hydrodynamic interactions in turbid media. The relevant polarization aspects on the physical quantities obtained from DWS experiments are detailed in Appendix K.

#### 4.4.2.2. Monte Carlo Simulations of Wave Propagation Through Three-Dimensional Inhomogeneous Media

The analysis in the previous Section 4.4.2.1 was confined to situations in which the size of the scatterers is very small relative to the wavelength of the probing radiation (i.e., Rayleigh region). We now focus on a natural extension of the

theory to cover the interesting situation of scattering by objects whose size is of the order of the wavelength or larger. We will use computational Monte Carlo algorithm to perform a simulation of the complete Stokes parameters for multiply scattered radiation in an inhomogeneous system composed of uncorrelated spherical particles. These simulations indicate that the amount of depolarization generated from multiple scattering depends on such factors as the size and shape distribution and index of refraction. The primary aim of this section is to present a theory for predicting the effect of particle size on the quantity of our primary interest—the degree of polarization—which is then compared with Monte Carlo simulation studies. These numerical results are compared to measurements on suspensions of polystyrene latex spheres in water.

The method of Monte Carlo simulation is well known in the context of statistical mechanics and condensed-matter physics; for a recent review, see Lewis and Miller [87]. The Monte Carlo modeling technique provides a way of finding solutions to multiple-scattering effects by tracing histories, that is, sequences of events that statistically occur to waves propagating through an optical medium. The following is a development of such a method. Here we are particularly concerned about the size parameter and optical depth dependences of the characteristic lengths of depolarization. A recent review article on Monte Carlo results that is complementary to the present section was presented by Bruscaglioni and colleagues [88]. The mathematical and statistical assumptions inherent in this procedure are well known in the literature and are not covered in depth here.

In this numerical experiment, one generates a realization of the random medium and calculates the resulting wavefield. We have used the SLAB Monte Carlo simulation code to analyze the depolarization behavior of a wave propagating through a slab of finite thickness and composed of uncorrelated spherical particles [47, 89]. This simulation technique was developed to study the three-dimensional random-walk-like multiple scattering process of the wave propagation. In the SLAB code, the three-dimensional paths for waves are followed from one scattering to the next as the wave propagates into the medium. Each scattering is assumed to be elastic and is described by the standard Mie theory. A single numerical simulation consists in launching some number of waves, at a source, along a specified axis. Referring to Fig. 4.4.1, a typical scattering path consists of a series of linear translations of random length (of average value equal to the mean free path  $l$ ), each of which is followed by a change of flight direction. Selection of the new flight direction is made by generating a random number from a scattering distribution function. The numerical implementation of this algorithm was checked through comparison with the Rayleigh regime for which exact analytic results are known. The theoretical details pertinent to the testing of the Monte Carlo code are reviewed in Ref. 89.

We still assume in our subsequent discussion the situation of weak scattering limit and that absorption can be regarded as negligible. Consider a

quasimonochromatic plane wavefield that is incident normally along the  $x_3$  axis on a plane parallel slab, of finite thickness  $d$  (where  $d \gg l$ ) and of infinite extent in the  $x_1, x_2$  directions, composed of uncorrelated spherical particles of radius  $a$  (Fig. 4.4.1). We begin by computing the degree of polarization of the light transmitted by the scattering medium for incident linearly ( $P_L$ ) and circularly ( $P_C$ ) polarized light. To do this, one must evaluate the different contributions of light following many different paths. Take a particular sequence of scattering events. The number of steps in this path of length  $s$  is  $n \equiv s/l$ . The number of scattering paths of length  $s$  is simply given by Green's function  $G(n, d)$  of the diffusion equation. The degrees of polarization are given by a proper weighting  $G(n, d)$  of scattering paths of length  $s$ . The resulting expressions are

$$P_i = \frac{\int_0^\infty f_i(n) G(n, d) dn}{\int_0^\infty G(n, d) dn} \quad (4.4.38)$$

where we have adopted the notation  $i = L$  for linear and  $C$  for circular states of polarization. In the multiple-scattering regime, the functions  $f$  express the dependences of the output degrees of polarization for a number of scatterings equal to  $n + 1$ . In the large  $n$  limit, the  $f$  values reduce simply to  $f_L(n) \cong \frac{3}{2} \exp[-n(l/\zeta_L)]$  and  $f_C(n) \cong \frac{3}{2} \exp[-n(l/\zeta_C)]$ . The  $\zeta$  terms define characteristic lengths of depolarization for a path of  $n + 1$  scatterings:  $\zeta_L = l/\ln(\frac{10}{7})$  and  $\zeta_C = l/\ln(2)$ . From the numerical values of  $\zeta$  we find that  $\zeta_L \cong 2\zeta_C$ . On performing the integration in (4.4.38), we obtain

$$P_i = \frac{d}{l} \frac{\sinh(l/\zeta_i)}{\sinh(d/\zeta_i)} \quad (4.4.39)$$

where  $\zeta_i = (\zeta_i l/3)^{1/2}$  with  $i = L, C$  define the characteristic lengths of depolarization for the slab geometry. Since  $d \gg \zeta_i$ , the degree of polarization of the transmitted light in the far field can be approximated by

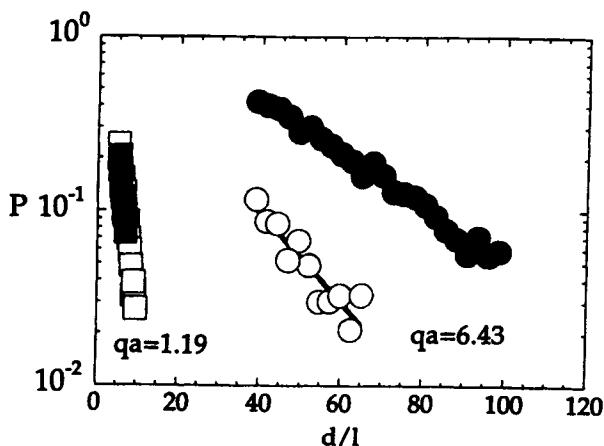
$$P_i \cong \frac{2d}{l} \sinh\left(\frac{l}{\zeta_i}\right) \exp\left(-\frac{d}{\zeta_i}\right) \quad (4.4.40)$$

Thus we see that the characteristic length of depolarization for incident linearly polarized light is greater (by a factor of  $\sqrt{2}$ ) than the corresponding length for incident circularly polarized light. This analysis should apply equally well for large spheres provided  $l$  is changed into the transport mean free path  $l^* = (1/\phi\sigma^*) = [l/[1 - \langle \cos(\theta) \rangle]]$ , where  $\langle \cos(\theta) \rangle$  is the mean cosine of the

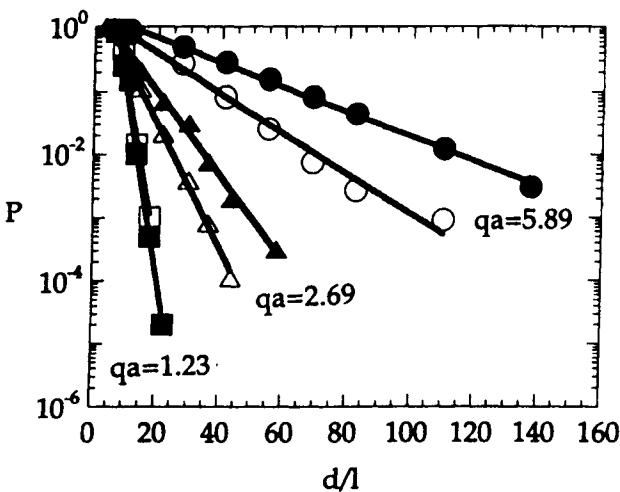
scattering angle  $\theta$ , and that the appropriate size dependence of the  $f$  values is inserted therein. Here the transport scattering cross section for each scatterer is defined in the usual way as  $\sigma^* = \int \sigma(\theta)(1 - \cos(\theta))d\Omega$ .

We now pass to the numerical results. Through the use of the numerical algorithm presented above, a set of different simulations was performed to investigate the effects of the particle size and medium thickness. The input parameters are the relative refractive index,  $m \equiv n_s/n_M = 1.20$ , where  $n_s$  and  $n_M$  are the refractive indices of the spheres ( $n_s = 1.59$  for polystyrene) and the surrounding medium ( $n_M = 1.33$  for water), the size parameter  $qa$ , and  $qI^* = 1000$ . These parameters were chosen for the purpose of comparison with experimental data. The experiments were carried out at room temperature, using a setup similar to that described in Ref. 90, which contains all relevant details. A semiconductor laser emitting at  $0.67 \mu\text{m}$  was used as the source beam. The beam was normally incident on one side of the sample (3 mm thickness), and the scattered light transmitted through the back wall of the sample cell was detected within a solid angle of collection of  $2^\circ$ . The scattering medium consists of various concentrations of polystyrene spheres (PolySciences, Inc.), with mean diameters of 0.22, 0.48, and  $1.05 \mu\text{m}$ , which were mixed into filtered distilled water and serve as the scattering centers. The mean cosines of the scattering angle were  $\langle \cos \theta \rangle = 0.342$  for spheres with mean diameters of  $0.22 \mu\text{m}$ ,  $\langle \cos \theta \rangle = 0.784$  for spheres with mean diameter of  $0.48 \mu\text{m}$ , and  $\langle \cos \theta \rangle = 0.917$  for spheres with mean diameters of  $1.05 \mu\text{m}$ .

Results of simulations that were carried out are presented in Figs. 4.4.6–4.4.9. For a starting point, we refer to Eq. (4.4.40) and find, in the limit  $qa \ll 1$ ,

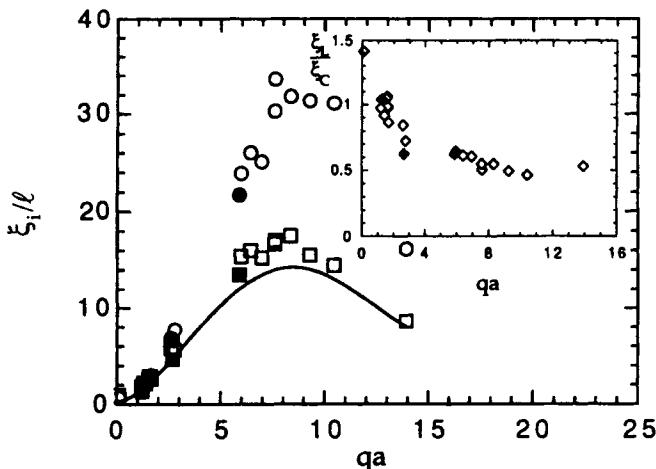


**Figure 4.4.6.** Semilogarithmic plot of the degrees of polarization for linearly polarized (open symbols) and circularly polarized (filled symbols) light as a function of the optical depth  $d/l$ . Squares represent the intermediate region  $qa = 1.19$ , and circles correspond to the Mie region  $qa = 6.43$ . The lines are exponential fits to the data.



**Figure 4.4.7.** Same as Fig. 4.4.6. Experimental data correspond to measurements on suspensions of polystyrene latex spheres in water ( $0.22\ \mu\text{m}$  ( $\circ$ )),  $0.48\ \mu\text{m}$  ( $\triangle$ ),  $1.05\ \mu\text{m}$  ( $\square$ )).

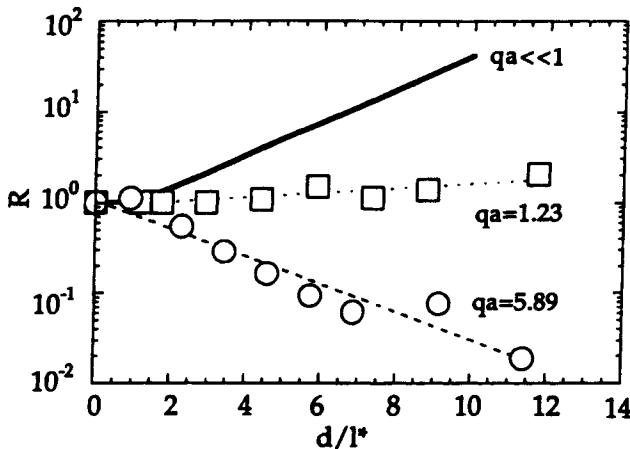
that  $\xi_L = 0.967 l$  and  $\xi_C = 0.684 l$ . The wave becomes depolarized over a distance that is of the order of the mean free path. Now examine Fig. 4.4.6, which shows the degrees of polarization for linearly and circularly polarized light as a function of  $d/l$  for  $qa = 1.19$  (i.e., intermediate Rayleigh–Mie region) and  $qa = 6.43$  (Mie region) in a semilogarithmic plot. The curves of Fig. 4.4.6 exhibit linear behavior in this plot. It is interesting to observe that the effect of the input polarization state in these regions is markedly different from that in the Rayleigh region. It is important to appreciate that for  $qa \sim 1$ , the slopes of these plots do not depend on the incident state of polarization. This is in contrast to the region  $qa > 1$ , for which these slopes ( $\equiv l/\xi_i$ ) now depend strongly on polarization; the slope is greater for linearly polarized light than for circularly polarized light. Having considered the numerics, we now proceed to a comparison with data from measurements on suspensions of polystyrene latex spheres. Variations of the degrees of polarization for three values of the size parameter (viz.,  $qa = 1.23$ ,  $2.69$ , and  $5.89$ ) are shown in Fig. 4.4.7. It is remarkable that the experiment gives an exponential decay over several decades. The behavior of these data is consistent with the simulation results of Fig. 4.4.6. Figure 4.4.8 shows the characteristic lengths of depolarization for incident linearly,  $\xi_L/l$ , and circularly,  $\xi_C/l$ , polarized light as a function of the dimensionless size parameter  $qa$ . In case of particles large compared to the wavelength, a linearly or circularly polarized wave becomes depolarized over a distance that is significantly greater than the mean free path. Both the experimental data and simulations show that the  $\xi$  values are nearly equal in



**Figure 4.4.8.** The characteristic lengths of depolarization,  $\xi_L/l$  for linearly polarized (o) and,  $\xi_C/l$ , for circularly polarized (O) light as a function of the dimensionless size parameter  $qa$ , as calculated by Monte Carlo simulation. Filled symbols indicate experimental data corresponding to measurements on suspensions of polystyrene latex spheres in water. Inset shows the ratio of the characteristic lengths  $\xi_L/\xi_C$  (open diamonds) as a function of  $qa$ . Experimental data (full diamonds) correspond to measurements on suspensions of polystyrene latex spheres in water. One sees that there is a good agreement with experiment. The solid curve shows  $l^*/l$  as a function of  $qa$ .

the region  $qa \sim 1$ , as can be seen in the inset of Fig. 4.4.8, where we have plotted the variation of the length ratio  $\xi_L/\xi_C$  with the size parameter  $qa$ . For the Rayleigh region, this ratio can be computed exactly using Eq. (4.4.40) and is equal to  $\sqrt{2}$ . In the range of sizes investigated, this ratio is a decreasing function of the parameter  $qa$ . As can be seen, the numerical calculations are in good agreement with experimental data. In Fig. 4.4.9 we have shown the variations of the ratio of the linear to circular degree of polarization  $R \equiv P_L/P_C$  as a function of  $d/l^*$ . We find that  $\ln(R) \sim d/l^*$ . Three behaviors are clearly seen in this plot. On one hand,  $R$  increases according to Eq. (4.4.40) in the Rayleigh regime. On the other hand, it can be seen that  $R$  is constant in the intermediate regime and decreases in the Mie regime.

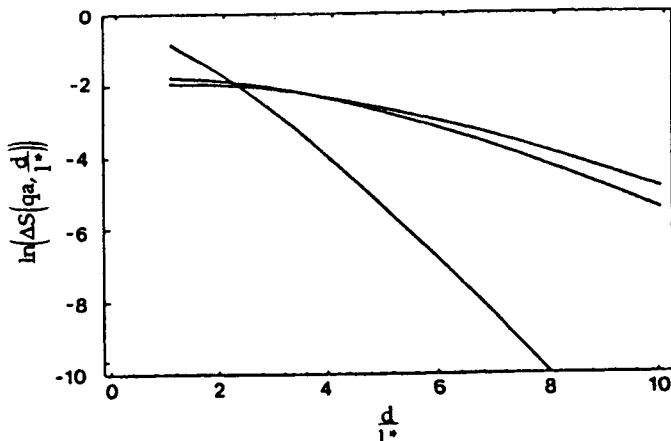
The difference between the two types of behavior corresponding to  $qa \ll 1$  and  $qa > 1$  stems from the anisotropy property of the scattering. In the Mie regime, the scattering is predominantly in the forward direction, while in the Rayleigh region, forward- and backward-scattered directions are treated on an equal footing. It is also appropriate to recall that the direction of linear polarization is not affected by a single scattering of light from a particle into the backward direction, regardless of particle size. It should be noted that the variations of  $\xi_L/l$  and  $l^*/l$  as function of  $qa$  are close to each other (i.e., solid



**Figure 4.4.9.** Semilogarithmic plot of the ratio of the degrees of polarization  $R \equiv P_L/P_C$  as a function of  $d/l^*$ , for three values of  $qa$ . Solid line corresponds to the Rayleigh region ( $qa \ll 1$ ). Symbols indicate experimental data corresponding to measurements on suspensions of polystyrene latex spheres in water [ $0.22\text{ }\mu\text{m}$  ( $\circ$ ),  $1.05\text{ }\mu\text{m}$  ( $\square$ )]. Dashed lines are fit to Eq. (4.4.40).

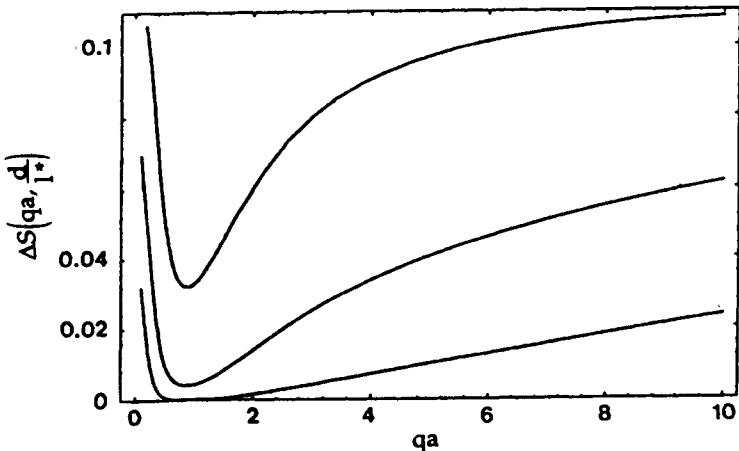
line in Fig. 4.4.8). This supports the idea that the mechanism of depolarization for an incident linearly polarized originates from the randomization of the direction. On the other hand, backscattering acts as an optical mirror (right  $\leftrightarrow$  left) for circularly polarized states, namely, helicity flip. The fact that the helicity is preserved over distances large compared to  $l^*$  for large scattering particles explains why the characteristic length for incident circularly polarized light is greater than the corresponding length for incident linearly polarized light. Two distinct mechanisms contribute to the depolarization of circularly polarized light: the randomization of the direction and the randomization of the helicity. However, it is difficult to infer from the preceding arguments what the contribution of these two mechanisms to the total length  $\xi_C$  will actually be. Two comments may be in order here. Since absorption is weak in most experiments, it has been neglected. However, we would like to stress that the effect of absorption may be an important issue in many cases. Moreover, it is now well established that when scatterers of a finite size are considered, Mie resonance effects (not considered here) should be taken into account [91]. In the resonance region of particle size parameters, the Rayleigh and geometric optics approximations are inapplicable, and numerical methods for characterizing scattering must be based on directly solving Maxwell's equation.

The discussion would be incomplete without considering some additional related developments concerning entropy production  $\Delta S \equiv S(qa, (d + l^*)/l^*) -$



**Figure 4.4.10.** Optical thickness, dependence of entropy production  $\Delta S \equiv S(qa, (d + l^*)/l^*) - S(qa, (d/l^*))$  plotted as  $\ln(\Delta S)$  versus  $d/l^*$ , for a circularly incident wave and fixed values of the size parameter  $qa$ . The values of  $qa$  from the top are 3.5, 5, and 7.

$S(qa, (d/l^*))$  [92]. The effect of optical thickness is illustrated in Fig. 4.4.10 for an incident circularly polarized wave. Since identical behavior is observed for incident linearly polarized waves, we concentrate on the circularly polarized case only. It is clear from looking at this figure that the entropy production falls off exponentially with the optical thickness for  $d/l^* \gg 1$ .



**Figure 4.4.11.** Entropy production as a function of the size parameter  $qa$  for a circularly polarized wave and fixed values of the optical thickness  $d/l^*$ . The values of  $d/l^*$  from the top are 10, 1, and 0.1.

The dependence of the entropy production on  $qa$ , hence on the degree of forwardness of the scattering, is depicted in Fig. 4.4.11. This figure shows a comparison between the size dependence of the entropy production for fixed values of  $d/l^*$ . Notice that a minimum of entropy production at  $qa \sim 1$ , corresponding to a maximum of entropy, separates a domain of decreasing entropy production for small particles from a domain of increasing entropy production for large particles. As before, this behavior is interpreted as arising from the anisotropy property of the scattering; furthermore, if we compare the three curves in Fig. 4.4.11, it is interesting to observe that  $\Delta S$  is significantly smaller when  $d/l^*$  increases.

At this point several comments are in order. The first of these is to clarify operationally what is meant by the "irreversible" evolution of the wave during scattering. Related to this, the second objective is to justify the introduction of the characteristic length scale of depolarization by a thermodynamic argument. The third objective is to discuss entropy production criteria for our specific situation.

Let us turn first to the physics behind irreversibility. The important problem of irreversibility in matter-radiation interactions has been the subject of some considerable debate. We first recall some results of an important paper of Clark Jones in which the emphasis was on obtaining a list of reversible and irreversible operations on a beam of light [93]. By means of heuristic arguments, Clark Jones assigned to a partially polarized beam two temperatures  $T_1$  and  $T_2$  and two values of entropy flux by the Planck's radiation formula [resp.  $S_1 = S(T_1)$  and  $S_2 = S(T_2)$ ]. When the beam is totally unpolarized (i.e.,  $T_1 = T_2 = T_u$ ), then  $S_1 = S_2 = S_u$ . On the other hand, for a completely polarized light (i.e.,  $T_1 \neq 0$ ,  $T_2 \neq 0$ ),  $S_1 \neq 0$  independent of the polarization form considered and  $S_2 = 0$ . The entropy becomes a function of macroscopic variables such as the radiance, frequency and temperature. Then, to quote Clark Jones:

The pleasant position has now been achieved where both a temperature and an entropy are associated with any given beam of light. It is now possible to decide the question of reversibility or irreversibility by thermodynamic criteria. According to the second principle of thermodynamics, an operation on a beam of light (supposed an isolated system) will be reversible if the total entropy is unchanged and irreversible if the total entropy increases natural processes in which the total entropy decreases are of course impossible.

The postulate of Jones is that the principles of macroscopic thermodynamic equilibrium are compatible with the description of the interaction of polarized light with any optical system. However, this simple thermodynamic analysis raised some criticisms as the following consideration shows. Let us consider the simple mixing of two completely polarized (one right circularly, indexed "r"; the other left circularly, indexed "l"): from the Jones point of view, the two components have temperature  $T_r = T_l = T_1$  and entropy  $S_1$ . As consequence of

the extensive character of entropy, we get for the mixture  $S = 2S_1$ , and  $T = T_1$ , which is not compatible with the fact that an incoherent superposition of two completely polarized beams can lead to a partially polarized light beam, even a completely unpolarized beam (see Appendix B for details). The key point to be stressed here is that any linear interaction is reversible if and only if  $S$  remains invariant under the transformation  $\mathbf{D}_2 \rightarrow \mathbf{D}_2$  and if no absorption occurs; otherwise it is irreversible. As a direct consequence, it is easy to prove that any unitary transformation  $\mathbf{D}_2 \rightarrow \mathbf{D}_2 = \mathbf{U}\mathbf{D}_2\mathbf{U}^{-1}$  constitutes a reversible operation on a beam of light, such as rotation of the state of polarization. From this point of view the problem of irreversibility can be introduced and defined when dissipation occurs in the medium (e.g. by selective absorption of states), or alternatively when taking into account correlations of phases and amplitudes of the field components induced by any interaction for which  $\Delta S \neq 0$ ; this could be achieved in practice by either a dilatating interaction ( $\Delta S < 0$ ; e.g., induced by a polarizer), or by a contracting depolarizing interaction ( $\Delta S > 0$ ; e.g., induced by a polarization scrambler).

The connection between entropy production and irreversibility has been discussed by many authors. At this stage, it remains to explain more precisely why a depolarization process actually implies that the system should tend to increase its entropy. The effect of multiple scattering on the propagation of waves is to randomize the incident wave vector direction, the phase of the electric field vector, and its polarization. Assume that the medium is static, which would be the case for a particular realization of the spatial disorder, and that the wave is quasimonochromatic with a coherence length that exceeds the characteristic path length  $d^2/l^*$ , namely, spectral bandwidth  $\Delta v \ll cl^{*2}/d^2$ . In such a case the time-invariant speckle pattern of scattered light, in which each speckle spot has a definite state of pure polarization, is characteristic of the given realization of disorder. Experimentally, Freund has shown that changing the polarization of the incident wave results in well-defined changes of the speckle pattern, and this can be used to reconstruct the Stokes vector of the incident beam by making speckle correlation measurements on the diffusively scattered light [94]. In this limit there is no entropy production, and the scattering process appears to be reversible. In the opposite situation,  $\Delta v \gg cl^{*2}/d^2$ , the speckle pattern is washed out and entropy can be produced. In like fashion, if the average is taken over all realizations of the disorder, such as by rotating the sample, the speckle pattern is washed out and entropy can be produced. In that situation, after undergoing a sufficiently large number of scattering events and ensemble-averaging the properties of the wave, the information contained in the polarization is lost. The irreversible radiative transfer from a low-polarization-temperature state to a high-polarization-temperature state results in a simple exponential law for the production of radiation entropy  $\Delta S \sim \exp[-(d/l)(l/\xi)]$ , where  $\xi$ , the characteristic length of depolarization, depends on whether the wave is initially linearly or circularly polarized and of the size of the spherical particles. The exponential law for the production of radiation entropy can be justified as follows. The production of radiation

entropy is given by

$$\Delta S = S\left(qa, \frac{d+l}{l}\right) - S\left(qa, \frac{d}{l}\right) = \ln\left(\frac{s(qa, d/l)}{s(qa, (d+l)/l)}\right) \quad (4.4.41)$$

This equation is further simplified by substituting the expansion  $\ln(s(x)) = \ln(2) + x^2/2 + O(x^4)$  for  $x \approx 0$  into it and making use of the definition of the correlation length  $\xi/l \sim |\ln(P)|^{-1}$  (Table 3.4.1); Eq. (4.4.41) can be rewritten as

$$\Delta S \sim \exp\left(-\frac{d}{\xi}\right) = \exp\left(-\left(\frac{d}{l}\right)\left(\frac{l}{\xi}\right)\right), \quad \frac{d}{l} \gg 1 \quad (4.4.42)$$

which establishes the result. It is remarkable that this exponential law is valid for both small-diameter ( $qa \ll 1$ ) and large-diameter ( $qa \gg 1$ ) spheres.

Up to this point we have illustrated a number of ways in which the entropy can be used to generate information related to multiple scattering of light by a spatially random medium. We should not leave this discussion without pointing out that radiation cannot reach a steady state without interaction with matter. This idea implies specific constraints on the entropy production. The principle of minimum entropy production was first discussed systematically by Prigogine [95], who argued, using ideas from the thermodynamics of irreversible processes (Onsager reciprocity theorem), that the entropy production reaches its absolute minimum rate in the steady state consistent with the constraints that prevent the system from reaching equilibrium. Reference is made to Prigogine [95] for an overview. Another entropy production minimum principle was put forward by Tykody, who proposed that the entropy production is a minimum in the steady state [96]. The situation discussed is a remarkable example where there are no conserved quantities other than the energy and for which the steady state is characterized by both maximum entropy and minimum production of entropy.

Regarding practical applications, these numerical and experimental results are relevant to the development of particle-sizing methods that are based on multiple scattering from laser beams and the characterization of Mie scatterers in turbid suspensions. In this chapter, the media studied with diffusing light have been isotropic. For additional literature on the theory of multiple scattering in anisotropic media, the reader may consult Ref. 97.

#### **4.4.2.3. Backscattering Enhancement from a Random Distribution of Scatterers**

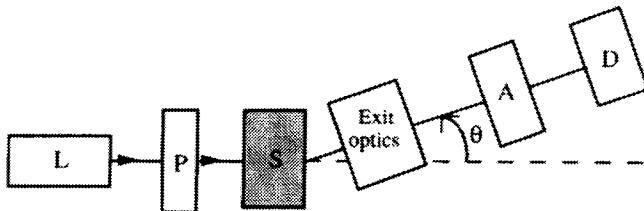
Enhanced backscattering of polarized light is one of the most remarkable phenomena associated with the scattering of light from a medium with randomly positioned scatterers [79, 80]. The important literature on this subject has shown that this phenomenon is recognized as a manifestation of

weak localization of waves. Since the effect of electron localization in metallic solids was proposed by Anderson in 1958 [98], it is well recognized that the localization effects have universal character and may be used to describe classic as well as quantum-mechanical disordered systems.<sup>10</sup> Hence they are a general phenomenon common to any wave propagation. The case of lightwaves is more interesting than the case of scalar waves because there are important effects associated with the polarization and the transverse nature of the lightwaves. For electrons in a disordered metal, the details of the scattering of the electrons by impurities is less accessible, and inelastic effects are generally more important. For a presentation of analogies between electron and optical wave phenomena, we refer the interested reader to Ref. [99]. The first observation of coherent backscattering of light from a disordered medium and its connection with electron localization was reported by van Albada and Lagendijk [100] and independently by Wolf and Maret [101].

The significant enhancement backscattering is observed in the form of a well-defined narrow peak in the angular distribution of the far-field intensity of the incoherent component of the scattered light at scattering angles near  $\pi$ . It has been argued that the effect results from the constructive interference between a wave following a certain multiple-scattering path and its conjugate wave following the time-reversed path [102]. The effect is most prominent when the particles are of the order of the wavelength in size and have single-scattering albedos. If pointlike particles are used as scatterers, the peak of enhanced intensity shows spatial anisotropy, as its width is greater in the direction of the polarization vector of the incident beam than in the perpendicular direction. The angular width of the peak is in the range of a few milliradians [103]. For an extensive review of experimental and theoretical works concerning the backscattering enhancement of the intensity of polarized light from volume scatterers such as particle suspensions, the reader is referred to Refs. [104] through [107]. There have been reports of similar phenomena concerning optically rough surfaces, either metallic or dielectric, whose standard deviation of surface height is several times larger than that of the illuminating wavelength [108] and random gratings [109, 110].

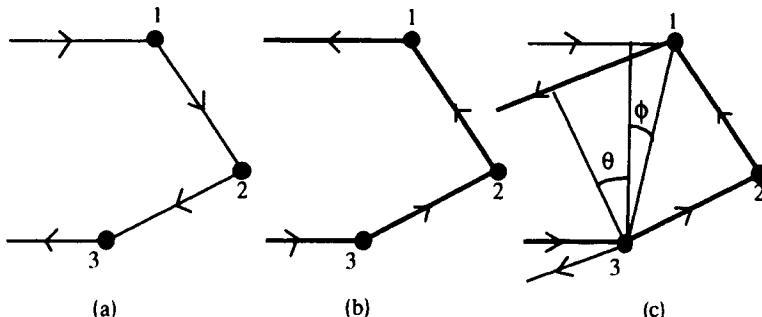
Consider the diffusing sample S made of a plane parallel medium composed of a suspension of identical uncorrelated dielectric spheres. The arrangement scheme to which the theory applies is depicted schematically in Fig. 4.4.12. The measuring principle for measuring the backscattering enhancement uses collimated light from a laser that is polarized (say, linearly) before entering a disordered dielectric slab. Light scattered at an angle  $\theta$  is detected by a photodetector mounted on a rotatable arm carrying the collecting optics and analyzer.

<sup>10</sup>Anderson had introduced the idea of localization, which occurs in a static disordered system, where electron interactions don't matter. Anderson localization originates from irregularities in a metal and cause localization of the wavefunction and render the metal insulating at low temperatures. As shown by Anderson [98], in a sufficiently disordered infinite material an entire band of electronic states can be spatially localized. The Anderson transition may be viewed as a transition from particlelike behavior described by the diffusion equation to wavelike behavior, which results in localization by interference.

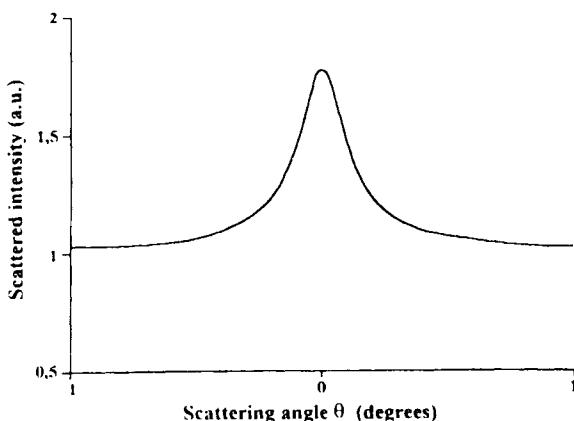


**Figure 4.4.12.** An operational diagram of the physical experiment used for measurement of the far-field angular dependence of the intensity of scattered light. The source beam is focused on the multiple-scattering sample  $S$ . The incident beam is polarized by means of the polarizer  $P$ . The intensity of the scattered light passed through an analyzer is measured by the photodetector  $D$ , which can be swung through an angle  $\theta = \pi$ .

We would like to give a heuristic explanation of the enhanced coherent backscattering phenomenon generated by a random set of particles. For that purpose we illustrate, in Fig. 4.4.13a, the multiple scattering of the incident beam by three scattering centers. To evaluate the intensity of the backscattered light, we need to add the amplitude from the conjugate wave (Fig. 4.4.13b). These two waves are in phase because they have equal path lengths and the backscattering intensity is doubled. In contrast, away from the backscattering direction by an angle  $\theta$ , the path difference  $\Delta$  between the wave and its conjugate can be written as  $\Delta = d(\sin(\theta + \phi) - \sin(\phi))$ , where  $d$  denotes the distance between the first and last scattering centers (Fig. 4.4.13c). We expect that the interference becomes unimportant when this is comparable to the wavelength  $\lambda$ . Consequently, it results in a doubling of the backscattering intensity in a narrow cone. A fundamental result is that for optically thick



**Figure 4.4.13.** Illustration of the basic phenomenon that underlies the enhanced backscattering of light as described in the text: (a) wave interacting with three scattering centers (sequence 123); (b) conjugate wave (sequence 321) which interferes constructively with that shown in (a); (c) scattering in a direction  $\theta$  away from  $\pi$  backscattering. Small values of  $\theta$  correspond to the domain of enhancement.



**Figure 4.4.14.** Intensity of backscattered light incident on a random collection of dielectric spheres, such as latex microspheres suspension in water as a function of scattering angle  $\theta$ . Note the enhancement by almost a factor of two in the backscattering directions.

media, the backscattering peak is characterized by a half-width at half maximum of the order of  $\lambda/l^* = \lambda/l(1 - \langle \cos(\theta) \rangle)$ .

By way of illustration, we present in Fig. 4.4.14 a typical curve of the shape of the backscattering cone formed under illumination of light for a random collection of dielectric spheres. Reference is made to Akkermans and colleagues for details of the experimental details [102]. The result shows clearly an on-axis enhancement in the average far-field irradiance distribution.

In conclusion, we note that the variety of physical phenomena brought to bear on the class of problems involving random media is a rich one, suggesting that a large amount of work still remains to be done in exploring this field. The next and final part will describe areas of recent intense interest and development in polarization research such as modeling of polarization of light propagating in anisotropic materials and measurement of polarization matrices of optical devices.

## FURTHER READING

1. E. L. O'Neill, *Introduction to Statistical Optics*, Addison-Wesley, Reading, MA, 1963.
2. A. Gerrard and J. M. Burch, *Introduction to Matrix Methods in Optics*, Wiley, London, 1975.
3. R. C. Jones, *J. Opt. Soc. Am.* **31**, 1941, 500; **32**, 1942, 486; **37**, 1947, 107; **38**, 1948, 671. See also Ref. 71 and H. Hurwitz and R. C. Jones, *J. Opt. Soc. Am.* **31**, 1941, 493.
4. H. Mueller, *J. Opt. Soc. Am.* **38**, 1948, 661.
5. J. D. Love and A. W. Snyder, *J. Opt. Soc. Am.* **65**, 1975, 1072.
6. Y. Fainman and J. S. Shamir, *Appl. Opt.* **23**, 1984, 3188.
7. N. G. Parke, *J. Math. Phys.* **28**, 1949, 131.
8. R. Barakat, *J. Opt. Soc. Am.* **53**, 1963, 317.
9. M. Born and E. Wolf, *Principles of Optics*, 6th ed., Pergamon Press, Oxford, 1980.
10. J. Byrne, *J. Phys. B: Atom. Molec. Phys.* **4**, 1971, 940.
11. A. Yariv and P. Yeh, *Optical Waves in Crystals*, Wiley, New York, 1984.
12. R. A. Chipman, *Opt. Eng.* **28**, 1989, 90.
13. F. Bretenaker and A. Le Floch, *J. Opt. Soc. Am.* **B8**, 1991, 230.
14. A. Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, New York, 1984.
15. C. Brosseau, *C. R. Acad. Sci.* **302**, Série II, **17**, 1986, 1047. See also C. Graham, *Proc. Roy. Soc. A* **369**, 1980, 517 and J. M. B. de Figueiredo and R. E. Raab, *Proc. Roy. Soc. A* **369**, 1980, 501.
16. W. A. Shurcliff, *Polarized Light: Production and Use*, Harvard Univ. Press, Cambridge, MA, 1966.
17. A. Jenkins and H. E. White, *Fundamentals of Optics*, McGraw-Hill, New York, 1950.
18. D. A. Holmes, *J. Opt. Soc. Am.* **54**, 1964, 1115.
19. C. Whitney, *J. Opt. Soc. Am.* **61**, 1971, 1207.
20. Y. Watanabe, *J. Appl. Phys.* **76**, 1994, 3994.
21. M. C. Pease, *Methods of Matrix Algebra*, Academic Press, New York, 1965.
22. I. Solc, *Cesk. Casopis Fys.* **3**, 1953, 366; *ibid.* **10**, 1960, 16; I. Solc, *Opt. Soc. Am.* **55**, 1965, 621. See also J. W. Evans, *J. Opt. Soc. Am.* **48**, 1958, 142.

23. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, New York, 1965.
24. R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light*, North-Holland, Amsterdam, 1977.
25. R. M. A. Azzam, *Appl. Phys.* **13**, 1977, 281.
26. C. Brosseau, *Optik* **88**, 1991, 109.
27. P. Soleillet, *Ann. Phys.* **12**, 1929, 23.
28. C. Brosseau and R. Barakat, *Opt. Commun.* **91**, 1992, 408.
29. C. V. M. van der Mee, *J. Math. Phys.* **34**, 1993, 5072.
30. A. B. Kostinski, *Appl. Opt.* **31**, 1992, 3506.
31. J. Cariou, B. Le Jeune, J. Lotrian and Y. Guern, *Appl. Opt.* **29**, 1990, 1689.
32. C. R. Givens and A. B. Kostinski, *J. Mod. Opt.* **41**, 1994, 471.
33. S. Y. Lu and R. A. Chipman, *J. Opt. Soc. Am.*, **11**, 1994, 766.
34. G. N. Ramachandran and S. Ramaseshan, "Crystal optics," in *Handbuch der Physik*, S. Flügge, ed., Springer-Verlag, Berlin, 1961, VI. XXV.
35. R. Barakat, *Opt. Commun.* **123**, 1996, 443.
36. H. Takenaka, *Nouv. Rev. Optique* **4**, 1973, 37. See also H. Takenaka, *Jpn. J. Appl. Phys.* **12**, 1973, 226, 1729.
37. E. P. Wigner, *Group Theory and Its Applications to the Quantum Mechanics of Atomic Spectra*, Academic Press, New York, 1959.
38. Z. F. Xing, *J. Mod. Opt.* **39**, 1992, 461.
39. These nine relations appear to have been derived explicitly for the first time by K. D. Abhyankar and A. L. Fymat, *J. Math. Phys.* **10**, 1969, 1935. See also E. S. Fry and G. W. Kattawar, *Appl. Opt.* **20**, 1981, 2811; J. W. Hovenier, *J. Atmos. Sci.* **26**, 1969, 488; R. Barakat, *Opt. Commun.* **38**, 1981, 159; J. W. Hovenier, H. C. van dee Hulst, and C. V. M. van der Mee, *Astron. Astrophys.* **157**, 1986, 301; L. Hecht, C. Hebsch, and B. Schrader, *Optik* **80**, 1988, 47; and I. Kuscer and M. Ribaric, *Optica Acta* **6**, 1959, 42.
40. J. Gil and E. Bernabeu, *Opt. Acta* **32**, 1985, 259.
41. C. Brosseau, *Optik* **88**, 1991, 109. See also C. Brosseau, *Optik* **85**, 1990, 180.
42. D. G. M. Anderson and R. Barakat, *J. Opt. Soc. Am.* **11**, 1994, 2305.
43. R. Barakat, *Opt. Commun.* **38**, 1981, 159.
44. P. Y. Gerligand, B. Le Jeune, J. Cariou, and J. Lotrian, *J. Phys. D: Appl. Phys.* **28**, 1995, 965.
45. C. Brosseau, *Opt. Commun.* **131**, 1996, 229.
46. D. A. Ramsey, Ph.D. dissertation, Univ. Michigan, Ann Arbor, 1985. See also M. W. Williams, *Appl. Opt.* **25**, 1986, 3616.
47. D. Bicout, C. Brosseau, A. S. Martinez, and J. M. Schmitt, *Phys. Rev. E* **49**, 1994, 1767.
48. D. W. Berreman, *J. Opt. Soc. Am.* **62**, 1972, 502. See also P. J. Lin-Ching and S. Teitler, *J. Opt. Soc. Am.* **A1**, 1984, 703 and S. Teitler and B. W. Henvis, *J. Opt. Soc. Am.* **60**, 1970, 830.
49. P. C. Waterman, *J. Appl. Phys.* **50**, 1979, 4550.

50. P. W. Barber and S. C. Hill, *Light Scattering by Particles: Computational Methods*, World Scientific, Singapore, 1990.
51. M. J. van Weert, *J. Opt. Soc. Am.* **68**, 1978, 1275.
52. P. Yeh, *J. Opt. Soc. Am.* **72**, 1982, 507.
53. C. Gu and P. Yeh, *J. Opt. Soc. Am. A* **10**, 1993, 966.
54. E. Collett, *Polarized Light: Fundamentals and Applications*, Marcel Dekker, New York, 1993.
55. R. M. A. Azzam, *Opt. Eng.* **20**, 1981, 58.
56. A. V. Rzhanov and K. K. Svitashhev, in *Advances in Electronics and Electron Physics*, Academic Press, New York, 1979, Vol. 49, p. 1.
57. F. Perrin, *J. Chem. Phys.* **10**, 1942, 415. See also F. Perrin and A. Abragam, *J. Phys. Rad.* **12**, 1951, 69.
58. R. S. Krishnan, *Proc. Indian Acad. Sci. A* **9**, 1938, 21; *ibid.* 1938, 91; *ibid.* 1938, 98. See also R. S. Krishnan, *Kolloid Zeits.* **84**, 1939, 2; *ibid.* **84**, 1938, 18; and R. S. Krishnan, *Proc. Indian Acad. Sci. A* **10**, 1939, 395.
59. S. Chandrasekhar, *Radiative Transfer*, Oxford Univ. Press, London, 1950.
60. H. C. van de Hulst, *Light Scattering by Small Particles*, Wiley, New York, 1957.
61. R. G. Newton, *Scattering Theory of Waves and Particles*, McGraw-Hill, New York, 1966.
62. J. I. Hage, J. M. Greenberg, and R. T. Wang, *Appl. Opt.* **30**, 1991, 1141.
63. S. Asano and M. Sato, *Appl. Opt.* **19**, 1980, 962. See also S. Asano and G. Yamamoto, *Appl. Opt.* **14**, 1975, 29.
64. E. M. Purcell and C. R. Pennypacker, *Astrophys. J.* **186**, 1973, 705. See also A. Lakhtakia, *Opt. Commun.* **79**, 1990, 1 for interesting details about the method and a comparison of the calculations with exact results for a sphere.
65. R. J. Perry, A. J. Hunt, and D. R. Huffman, *Appl. Opt.* **17**, 1978, 2700.
66. C. R. Hu, G. W. Kattawar, M. E. Parkin, and P. Herb, *Appl. Opt.* **26**, 1987, 4159.
67. Y. Shi and W. McClain, *J. Chem. Phys.* **98**, 1993, 1695.
68. J. R. Wait, *Appl. Sci. Res.* **10**, 1963, 441.
69. S. B. Singham and G. C. Salzman, *J. Chem. Phys.* **84**, 1986, 2658. See also M. K. Singham, S. B. Singham, and G. C. Salzman, *J. Chem. Phys.* **85**, 1986, 3807.
70. C. Brosseau, *Pure Appl. Opt.* **5**, 1996, 755.
71. D. S. Saxon, *Phys. Rev.* **100**, 1955, 1771.
72. Z. Sekera, *J. Opt. Soc. Am.* **56**, 1966, 1732.
73. R. M. A. Azzam, *Optica Acta* **24**, 1977, 1039.
74. N. Vansteenkiste, P. Vignolo, and A. Aspect, *J. Opt. Soc. Am. A* **10**, 1993, 2240.
75. M. Martinelli, *Opt. Commun.* **72**, 1989, 341.
76. R. Bhandari, *Opt. Commun.* **88**, 1992, 1.
77. A. Schönhofer and H. G. Kuball, *Chem. Phys.* **115**, 1987, 159.
78. A. Ishimaru, *Wave Propagation and Scattering in Random Media*, Academic Press, New York, 1978.
79. See the feature issue on wave propagation and scattering in random media, *J. Opt. Soc. Am. A* **2**, 1985, 206 and references cited therein.

80. P. Sheng, ed., *Scattering and Localization of Classical Waves in Random Media*, World Scientific, Singapore, 1990.
81. M. Nieto-Vesperinas and J. C. Dainty, eds., *Scattering in Volumes and Surfaces*, North-Holland, Amsterdam, 1990.
82. Y. N. Barabanenkov, A. Kravtsov, V. D. Ozrin, and A. I. Saichev, "Enhanced backscattering in optics," in *Progress in Optics*, E. Wolf, ed., Elsevier, New York, 1991, Vol. 29.
83. C. Brosseau and R. Barakat, *Opt. Commun.* **84**, 1991, 127.
84. K. Kim, L. Mandel, and E. Wolf, *J. Opt. Soc. Am. A4*, 1987, 433.
85. D. Bicout and C. Brosseau, *J. Phys. I France* **2**, 1992, 2047. See also I. Freund, *Opt. Commun.* **81**, 1991, 251 and E. Akkermans, Ph.D. thesis, Fourier Univ., Grenoble, France, 1986.
86. C. Brosseau, *Appl. Opt.* **34**, 1995, 4788.
87. E. E. Lewis and W. F. Miller, *Computational Methods of Neutron Transport*, Wiley, New York, 1984.
88. P. Bruscaglioni, G. Zaccanti, and Q. Wei, *Appl. Opt.* **32**, 1993, 6142. See also P. Bruscaglioni and G. Zaccanti, in *Multiple Scattering in Dense Media*, M. Nieto-Vesperinas and J. C. Dainty, eds., Elsevier, New York, 1990.
89. A. S. Martinez, Ph.D. dissertation, Fourier Univ., Grenoble, France, 1993.
90. J. M. Schmitt, A. H. Gandjbakhche, and R. F. Bonner, *Appl. Opt.* **31**, 1992, 6535.
91. C. F. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles*, Wiley, New York, 1983.
92. C. Brosseau and D. Bicout, *Phys. Rev. E* **50**, 1994, 4997.
93. R. C. Jones, *J. Opt. Soc. Am.* **43**, 1953, 138.
94. I. Freund, *Opt. Lett.* **15**, 1990, 1425. Similarly, for diffusing (Brownian motion) scatterers, if the speckle pattern is studied on a timescale short compared to the multiple-scattering dephasing time  $\tau \equiv (cl^*{}^2/\omega d^2D)$ , where  $D$  is the particle diffusion constant, then a well-defined speckle pattern is again obtained with no entropy production; see for instance, M. Rosenbluh, M. Hoshen, I. Freund, and M. Kaveh, *Phys. Rev. Lett.* **58**, 1987, 2754.
95. I. Prigogine, *Thermodynamics of Irreversible Processes*, Wiley, New York, 1971.
96. R. Tykody, *Thermodynamics of Steady States*, Macmillan, New York, 1967.
97. M. I. Mischenko, *J. Opt. Soc. Am. A8*, 1991, 871 and M. I. Mischenko, *J. Opt. Soc. Am. A9*, 1992, 978. See also the feature issue "Diffusing Photons in Turbid Media," *J. Opt. Soc. Am. A14*, 1997, 136.
98. P. W. Anderson, *Phys. Rev.* **109**, 1958, 1492.
99. S. John, *Phys. Today*, **1991**, 32.
100. M. P. van Albada and A. Lagendijk, *Phys. Rev. Lett.* **55**, 1985, 2692.
101. P. E. Wolf and G. Maret, *Phys. Rev. Lett.* **55**, 1985, 2696.
102. E. Akkermans, P. E. Wolf, R. Maynard, and G. Maret, *J. Phys. (Paris)* **49**, 1988, 77. See also P. E. Wolf, G. Maret, E. Akkermans, and R. Maynard, *J. Phys. (Paris)* **49**, 1988, 63; E. Akkermans, P. E. Wolf, and R. Maynard, *Phys. Rev. Lett.* **56**, 1986, 1471; M. J. Stephen and G. Cwilich, *Phys. Rev. B* **34**, 1986, 7564 and M. J. Stephen, *Phys. Rev. B* **37**, 1988, 1.

103. M. P. van Albada, M. B. van der Mark, and A. Lagendijk, *Phys. Rev. Lett.* **58**, 1987, 361.
104. L. Tsang and A. Ishimaru, *J. Opt. Soc. Am. A1*, 1984, 831. See also A. Ishimaru and L. Tsang, *J. Opt. Soc. Am. A5*, 1988, 228, Y. Kuga, L. Tsang, and A. Ishimaru, *J. Opt. Soc. Am. A2*, 1985, 616, and C. E. Mandt, L. Tsang, and A. Ishimaru, *J. Opt. Soc. Am. A7*, 1990, 585.
105. A. A. Golubentzev, *Zh. Eksp. Teor. Fiz.* **86**, 1984, 47 (*Sov. Phys. JETP* **59**, 1984, 26).
106. A. Ishimaru, *Wave Propagation and Scattering in Random Media*, Academic Press, New York, 1978.
107. Yu. A. Kravtsov, S. M. Rylov, and V. I. Tatarskii, *Usp. Fiz. Nauk* **115**, 1975, 239 (*Sov. Phys. Usp.* **18**, 1975, 118).
108. K. A. O'Donnell and E. R. Mendez, *J. Opt. Soc. Am. A4*, 1987, 1184.
109. A. A. Maradudin, T. Michel, A. R. McGurn, and E. R. Mendez, *Ann. Phys.* **203**, 1990, 255.
110. M. Nieto-Vesperinas and J. M. Soto-Crespo, *Phys. Rev. B38*, 1988, 7250; *ibid.* 8193.

## PART 5

---

# APPLICATIONS TO SELECTED TOPICS

Having discussed fundamental aspects of polarization optics, we will now develop applications of the formalisms in somewhat more concrete terms. The spread of polarization applications of light from optical imaging to fiber sensors to optical data processing and optical communications is remarkable. This is by no means an exhaustive list. The subject of polarization is diverse and interdisciplinary, and it is rapidly evolving. Chemists, physicists, material scientists, and engineers are all active contributors. On the technological side, results are being consolidated, mainly because of a considerable improvement in instrumentation. It is a mere fact that significant advances in science take place when the state of the art in instrumentation improves. For example, new physical effects have been identified in elastic multiple light scattering by ensembles of particles that are relevant to the imaging of tissue structure with diffusing light [1]. The development of optical technologies requires an understanding of the physics behind the phenomena of propagation and scattering of partially polarized light in these media on a fundamental level.

A general optical device is a complicated polarization modification system containing numerous internal reflections, birefringent elements, and so on that allow for partial polarizing or depolarizing. A great many sophisticated instruments and techniques, developed largely to characterize light polarization, such as polarimeters and ellipsometers, have been the subject of repeated attention over several decades. Their potential as polarization devices has been recognized for just as long; they are technologically mature, and numerous references review their operating characteristics. The polarization analysis of these optical devices has two aspects. First, there is an engineering problem—it is quite difficult to design optical systems that leads to precise control of polarization. This issue, a delicate optimization problem in many variables, is better left to the specialists. Second, a thorough understanding of polarization transformation and depolarization in these systems must be characterized by basic physics through a discussion in terms of Jones and Mueller polarization matrices. This final part will mainly concern this second issue. Three remarks are appropriate at this point. First we note that the subject is much too broad and includes too wide a range of diverse applications to receive a detailed treatment. It is no wonder that we sometimes feel overwhelmed by information. No single author can master this vast field and summarize the current state of

the knowledge. We have made the choice of giving a detailed exposition of a limited number of carefully selected examples rather than providing a wider but more superficial discussion of more topics. These omissions will be in addition to those we are already guilty of by either choice or default. We used our own judgment and prejudices to select the subjects that we considered more relevant so that the reader can see the theoretical framework developed in the previous chapters at work. The coverage of these topics is sufficient to gain a first-order understanding of the subject explained in physical terms. Second on the experimental side, the discussions are brief. Third, a number of topics have been omitted, such as polarization aberration and ray tracing, polarimetry for radar remote sensing. Also we do not try to recapitulate all the data on a given problem but rather select studies that give a relevant overview of the state of art. Selected up-to-date references provide opportunities for the reader to look for further material.

The material is divided into four main sections. In Section 5.1, we address the issue of change of polarization of a plane electromagnetic wave propagating through anisotropic materials. This interaction may induce a transformation as well as a depolarization of the incident wave. After introducing differential polarization matrices, we develop some useful relations concerning the behavior of the ellipse of polarization in media, where optical parameters depend on position along the direction of propagation of light. An alternative description involving the Stokes parameters is discussed, which permits us to describe the propagation of partially polarized light in depolarizing anisotropic media. As illustrative examples, we consider a brief review of the physics behind the propagation of partially polarized light in liquid crystals. Here we derive the specific expressions related to the evolution of the Stokes parameters in these media. Section 5.2 is devoted to a brief description of the usual optical polarizing components (polarizer, compensator, rotator) that modify the state of polarization of incident illumination. Sections 5.3 and 5.4 cover the aspects of measurements of the Stokes parameters and the polarization matrices by well-chosen experiments. Here the important question deals with establishing the information content of the various polarization matrix elements.

## SECTION 5.1

---

# Electromagnetic Propagation in Linear Anisotropic Media

The purpose of this section is to describe in some detail the formalisms of relevance to characterize the electromagnetic propagation in linear anisotropic media. Optically anisotropic media are materials whose optical properties naturally depend on the direction of propagation of the light (e.g., calcite, liquid crystals), or are induced by some external source (e.g., electrooptic effect), or stress (e.g., subsequent to mechanical deformation, polymeric materials exhibit anisotropic optical properties as a result of microstructural orientation). Two basic optical properties are of importance: birefringence—the anisotropic retardation of the light and dichroism—the anisotropic attenuation of the light. For example, quartz is a birefringent crystal and tourmaline (boron silicate) is a dichroic crystal.

To retain the discussion within bounds, we concentrate mainly on the polarization dependence on the propagation of electromagnetic waves in anisotropic media. It is the purpose of Section 5.1.1 to present a formalism for describing the evolution of polarization into a linear nonmagnetic optical material having anisotropic permittivity. Our derivation of the basic equations that describe the propagation of light in a linear medium with permittivity tensor start with Maxwell's equations. A system of differential equations that describe how the Jones and Stokes vectors change as a wave propagates through a linear anisotropic layer between two parallel planes orthogonal to the direction of propagation of light is considered in Section 5.1.2. More specifically, we introduce and discuss Jones and Mueller differential polarization matrices that provide a simple and convenient way to deal with the description of wave propagation in anisotropic materials. Furthermore, the lamellae theory of Clark Jones and the matrix exponential function that led to the identification of eight types of elementary behavior associated with birefringence and dichroism are reviewed. Next, Section 5.1.3 deals with the complex polarization ratio representation of polarized light. The evolution of the state of polarization under the action of birefringence and dichroism is described by way of the Stokes-vector space representation of polarization in Section 5.1.4. We mention the importance of these formalisms to the fields of liquid crystals in Section 5.1.5. For alternate discussions on the general subject of wave propagation in anisotropic media, the reader may wish to consult Refs. [1–6].

### 5.1.1. PLANE-WAVE PROPAGATION IN A LINEAR MEDIUM WITH PERMITTIVITY TENSOR

Before we further develop the ideas of differential polarization matrices, let us generalize the formalism of Part 2 to permit it to handle the propagation of a plane wave in a linear anisotropic medium. The sample material is completely general; it could be a crystal or a fluid subject to applied fields. For simplicity, we assume that the sample is usually in a charge-free and current-free environment, that is, that in the sense of the description of electromagnetism, there are no external sources. To accomplish this, we return briefly to Section 2.1 of Part 2 and specify the notation that we shall use in the rest of this chapter. The theory presented in Section 2.1 concerned linear, isotropic, homogeneous, and nonabsorbing media, with their permittivity and permeability represented by real scalars. In isotropic media the electric displacement vector is related to the electric field vector by a scalar quantity: the permittivity. The square root of the permittivity is the refractive index for the waves propagating in the medium.

Next we will consider an optical field in the form of quasimonochromatic plane waves incident on a homogeneous, nonabsorbing, nonmagnetic material with permeability equal to  $\mu_0$  and having anisotropy in permittivity. We see from the material equation Eq. (2.1.2a) that  $\mathbf{D}$  is not generally parallel to  $\mathbf{E}$  as for an isotropic medium. The anisotropy is described by a second-order permittivity tensor

$$\boldsymbol{\epsilon} = \epsilon_0 \begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{21} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{31} & \epsilon_{32} & \epsilon_{33} \end{bmatrix} \quad (5.1.1)$$

which can be allowed to have nine independent nonzero elements. Restricting our consideration to non-optically active dielectric media,  $\mathbf{E}$  and  $\mathbf{D}$  are related by a real tensor. For a lossless optical medium in which magnetooptic effects are absent, the dielectric tensor is symmetric and has, in general, only six independent elements. A point that deserves to be mentioned is that conservation of electromagnetic field energy requires that  $\epsilon_{ij} = \epsilon_{ji}$ , in other words, that the permittivity tensor be symmetric. The proof follows from Poynting's theorem [Eq. (2.3.27)] [4]. Thus the permittivity tensor requires at most six independent parameters to parametrize it. In point of fact, the number of independent parameters is under the dependence of the type of crystal considered (see Table 5.1.1).

The dielectric tensor for a medium oriented with its principal axes along the axes of a coordinate system reduces to a diagonal representation

$$\boldsymbol{\epsilon} = \epsilon_0 \begin{bmatrix} \epsilon_1 & 0 & 0 \\ 0 & \epsilon_2 & 0 \\ 0 & 0 & \epsilon_3 \end{bmatrix} \quad (5.1.2)$$

**TABLE 5.1.1. Classification of Anisotropic Crystals**

Optical Symmetry	Type of Crystal	Dielectric Tensor $\epsilon$
Isotropic	Cubic	$\epsilon = \epsilon_0 n^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$
Uniaxial	Tetragonal, hexagonal, trigonal	$\epsilon = \epsilon_0 \begin{bmatrix} n_o^2 & 0 & 0 \\ 0 & n_o^2 & 0 \\ 0 & 0 & n_e^2 \end{bmatrix}$
Biaxial	Triclinic, monoclinic, orthorhombic	$\epsilon = \epsilon_0 \begin{bmatrix} n_1^2 & 0 & 0 \\ 0 & n_2^2 & 0 \\ 0 & 0 & n_3^2 \end{bmatrix}$

and the  $n_k = (\epsilon_k)^{1/2}$  ( $k = 1, 2, 3$ ) define the principal indices of refraction. These parameters can be interpreted as lengths and angular orientations of the principal axes with respect to the crystallographic axes [4]. The usual classification of anisotropic crystals is as follows: isotropic materials are such that  $n_1 = n_2 = n_3 = n$ . Only one parameter, the refractive index, is needed to describe the dielectric tensor. Uniaxial materials are such  $n_1 = n_2 = n_o \neq n_3 = n_e$  with the  $x_3$ -axis defining the optic axis. Two parameters are needed to describe the dielectric tensor. Unpolarized light incident on a uniaxial crystal produces two linearly polarized beams: the ordinary and the extraordinary beams with respective refractive indices  $n_o$ , which is a constant and independent of direction; and  $n_e$  which changes with direction. These two indices of refraction are equal only in the direction of the optic axis. Biaxial materials are such that  $n_1 \neq n_2 \neq n_3$ . The numerical difference  $\Delta n = n_e - n_o$  is a measure of the birefringence; it can be positive or negative. Calcite is an example of negative uniaxial crystal. Quartz is a positive uniaxial crystal. Both  $n_o$  and  $n_e$  in these materials vary with propagation direction. We shall return soon to the specificity of the ordinary and extraordinary indices. The optical symmetries of the crystals along with their corresponding dielectric tensor are listed in Table 5.1.1.

We turn now to the main point of this section and characterize the plane-wave propagation in an anisotropic lossless medium. We begin by assuming spatial and temporal dependences of the vector fields  $\mathbf{E}$  and  $\mathbf{H}$  of the form  $\exp(-i(2\pi vt - q\mathbf{e}_3 \cdot \mathbf{R}))$ , where  $\mathbf{q} = q\mathbf{e}_3$  is the wave vector. Returning to Maxwell's equations in complex form [Eqs. (2.3.5a-d)], we obtain

$$\mathbf{e}_3 \cdot \mathbf{D} = 0 \quad (5.1.3a)$$

$$\mathbf{e}_3 \cdot \mathbf{B} = 0 \quad (5.1.3b)$$

$$\nabla \times \mathbf{E} = 2\pi i v \mu_0 \mathbf{H} = i\mathbf{q} \times \mathbf{E} \quad (5.1.3c)$$

$$\nabla \times \mathbf{H} = -2\pi i v \mathbf{D} = i\mathbf{q} \times \mathbf{H} \quad (5.1.3d)$$

It is important to appreciate that  $\mathbf{D}$  and  $\mathbf{H}$  are perpendicular to  $\mathbf{q}$ , but not  $\mathbf{E}$  which is normal to the Poynting vector  $\mathbf{P}$ .

By further eliminating  $\mathbf{H}$  from Eqs. (5.1.3c,d), we arrive at the relation

$$4\pi^2 v^2 \mu_0 \mathbf{D} = q^2 (\mathbf{E} - (\mathbf{E} \cdot \mathbf{q}) \mathbf{q}) \quad (5.1.4)$$

if we make use of the triple cross-product vector identity [7]

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = (\mathbf{a} \cdot \mathbf{c}) \mathbf{b} - (\mathbf{a} \cdot \mathbf{b}) \mathbf{c} \quad (5.1.5)$$

where  $\mathbf{a}$ ,  $\mathbf{b}$  and  $\mathbf{c}$  are three arbitrary vectors. Equation (5.1.4) can also be rewritten as

$$\epsilon \mu_0 \mathbf{E} = \frac{q^2}{4\pi^2 v^2} (\mathbf{E} - (\mathbf{E} \cdot \mathbf{q}) \mathbf{q}) \quad (5.1.6)$$

or equivalently in terms of components  $E_k$

$$\epsilon_0 \epsilon_k \mu_0 E_k = \frac{q^2}{4\pi^2 v^2} (E_k - (\mathbf{E} \cdot \mathbf{q}) q_k) \quad (5.1.7)$$

where  $\epsilon_k$  are the components of the dielectric tensor in the principal coordinate system. Now define, in a manner similar to Eq. (2.3.11), the phase velocity

$$v = 2\pi \frac{v}{|\mathbf{q}|}$$

for the plane wave and the velocity  $v_k = 1/(\epsilon_0 \mu_0 \epsilon_k)^{1/2} = c/(\epsilon_k)^{1/2} = c/n_k$  in the direction characterized by index  $k = 1, 2, 3$ . In terms of these definitions, Eq. (5.1.7) reads as

$$E_k = \frac{v_k^2 (\mathbf{E} \cdot \mathbf{q}) q_k}{v_k^2 - v^2} \quad (5.1.8)$$

As the reader can determine, the  $E_k$  can be eliminated, and we obtain the expression

$$\sum_{k=1}^3 \frac{q_k^2}{v^2 - v_k^2} = 0 \quad (5.1.9)$$

which was first derived by Fresnel and for that reason, is termed the *Fresnel equation of wave normals*. Equation (5.1.9) provides the associated velocities with propagation direction  $\mathbf{q}$ . It is worth noting that in  $\mathbf{q}$  space, Eq. (5.1.9) can be represented by a three-dimensional surface.<sup>1</sup> As a final note, it is worth emphasizing that Eq. (5.1.9) is quadratic in  $v^2$ . Therefore, for each direction of

<sup>1</sup>An ellipsoid in Cartesian coordinates is defined as  $(x_1/a_1)^2 + (x_2/a_2)^2 + (x_3/a_3)^2 = 1$  in which the  $x_i$  ( $i = 1, 2, 3$ ) are the symmetry axes of the ellipsoid with semiaxes of values  $(a_j)^{1/2}$ .

propagation,  $v^2$  takes two possible values. Consequently the medium has two linearly polarized beams with two different phase velocities, and the two electric displacement vectors are mutually orthogonal. A proof of this important result is outlined below. Let  $\mathbf{E}_1$  and  $\mathbf{E}_2$  be the electric field vectors and  $\mathbf{D}_1$  and  $\mathbf{D}_2$  be the displacement vectors of these two beams. Using Eqs. (5.1.3a) and (5.1.4) we have

$$\mathbf{D}_1 \cdot \mathbf{D}_2 = 0 \quad (5.1.10a)$$

$$\mathbf{D}_1 \cdot \mathbf{E}_2 = 0 \quad (5.1.10b)$$

$$\mathbf{D}_2 \cdot \mathbf{E}_1 = 0 \quad (5.1.10c)$$

Observe that  $\mathbf{E}_1$  and  $\mathbf{E}_2$  are, in general, not orthogonal. As a final remark, we may write the orthogonality relation

$$\mathbf{e}_3 \cdot (\mathbf{E}_1 \times \mathbf{H}_2) = \mathbf{e}_3 \cdot (\mathbf{E}_2 \times \mathbf{H}_1) = 0 \quad (5.1.11)$$

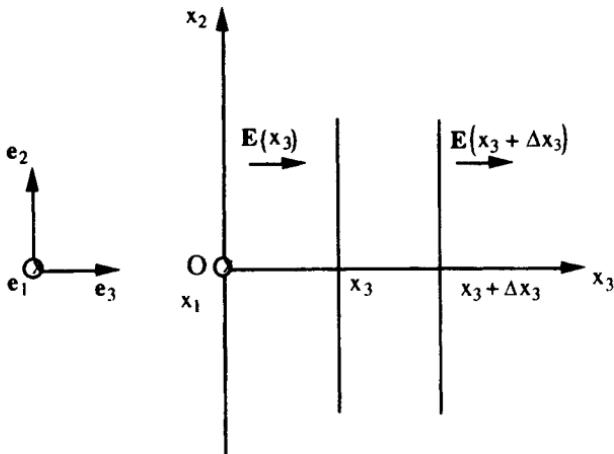
This expression is important because it shows that the power flow in an anisotropic medium along the direction of propagation is the sum of the power carried by each beam individually [4]. In summary, the two displacement vectors  $\mathbf{D}_1$  and  $\mathbf{D}_2$ , corresponding to the two normal modes of propagation, have well-defined polarization states and phase velocities.

With this material on wave propagation in anisotropic media behind us, we can now focus on the differential polarization matrices.

### 5.1.2. THE DIFFERENTIAL POLARIZATION MATRICES

Having characterized the effect of optical elements of finite thickness, we now consider the propagation of the Jones and Stokes vectors of the beam that propagates through linear anisotropic media as a function of distance into the medium, by means of a differential equation.

Consider first the case of a deterministic optical medium that cannot decrease the degree of polarization of an incident pure state of polarization. As I discussed in Part 4, the optical properties of such a medium can be described by the Jones matrix  $\mathbf{J}$ . Now we consider the propagation of the wave through the medium as a function of distance into the medium. The approach used here starts by assuming that the state of polarization of the wave changes continuously as it propagates through an optically anisotropic deterministic medium that is inhomogeneous along the  $x_3$  direction that is perpendicular to the interface. The basic geometry of the situation is outlined in Fig. 5.1.1. Let the optical medium be transversally homogeneous within any plane perpendicular to the light path. The incident optical field propagating as a beam (paraxial approach) in the direction of a unit vector  $\mathbf{e}_3$  is a function of coordinate  $x_3$  only. Only forward-propagating plane waves are being considered. Assuming



**Figure 5.1.1.** Illustration of notation related to the propagation of light in a slab of a linear anisotropic layer bounded by the planes parallel to the  $(x_1, x_2)$  plane located at  $x_3$  and the plane parallel to it a distance  $\Delta x_3$  away along the direction of propagation;  $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$  is a right-handed orthogonal basis;  $x_3 = 0$  denotes the origin of the optical medium; and  $\mathbf{E}(x_3)$  and  $\mathbf{E}(x_3 + \Delta x_3)$  are the Jones vectors of the wave at two transverse planes located at  $x_3$  and  $x_3 + \Delta x_3$ , respectively.

quasimonochromatic light, it follows immediately from Eq. (4.1.1) that

$$\mathbf{E}(x_3 + \Delta x_3) = \mathbf{J}(x_3, \Delta x_3)\mathbf{E}(x_3) \quad (5.1.12)$$

where  $\mathbf{J}(x_3, \Delta x_3)$  denotes the Jones matrix of a thin slab of the medium bounded at  $x_3$  and  $x_3 + \Delta x_3$  by two planes perpendicular to the  $x_3$  axis.<sup>2</sup> Subtracting  $\mathbf{E}(x_3)$  on both sides of Eq. (5.1.12), we find that

$$\mathbf{E}(x_3 + \Delta x_3) - \mathbf{E}(x_3) = (\mathbf{J}(x_3, \Delta x_3) - \boldsymbol{\sigma}_0)\mathbf{E}(x_3) \quad (5.1.13)$$

The optical properties of a thin layer between the surfaces having the coordinates  $x_3$  and  $x_3 + \Delta x_3$ , with both two plane boundaries parallel to the  $(\mathbf{e}_1, \mathbf{e}_2)$  plane (Fig. 5.1.1), is described by the differential Jones matrix [8]

$$\mathbf{j}(x_3) \equiv \lim_{\Delta x_3 \rightarrow 0} \frac{\mathbf{J}(x_3, \Delta x_3) - \boldsymbol{\sigma}_0}{\Delta x_3} \quad (5.1.14)$$

It is worth observing that the elements of the  $\mathbf{j}$  matrix have the dimension of

<sup>2</sup>A note about terminology. Clark Jones has used the electric field vector  $\mathbf{E}$  as a fundamental variable of his theory. In anisotropic optical media, it is the electric displacement vector  $\mathbf{D}$  that should represent the optical field. However, we shall adopt the prevailing usage in the literature of  $\mathbf{E}$  rather than  $\mathbf{D}$  vectors.

the inverse of a length. Let us illustrate by an example—for the purpose of this example I assume that each layer is a homogeneous linearly birefringent and nonabsorbing plate, the principal axes of which are parallel to the  $x_1$  and  $x_2$  directions. From Eq. (4.1.35), we recognize that

$$\begin{aligned} \mathbf{J}(x_3, \Delta x_3) &= \begin{bmatrix} \exp\left(-\frac{2i\pi v n_1}{c} \Delta x_3\right) & 0 \\ 0 & \exp\left(-\frac{2i\pi v n_2}{c} \Delta x_3\right) \end{bmatrix} \\ &\cong \begin{bmatrix} 1 - \frac{2i\pi v n_1}{c} \Delta x_3 & 0 \\ 0 & 1 - \frac{2i\pi v n_2}{c} \Delta x_3 \end{bmatrix} \end{aligned} \quad (5.1.15)$$

where  $n_1$  and  $n_2$  characterize the birefringence along the  $x_1$  and  $x_2$  directions, respectively. We see that relation (5.1.14) implies that

$$\mathbf{j}(x_3) = \begin{bmatrix} -2i\frac{\pi v}{c} n_1 & 0 \\ 0 & -2i\frac{\pi v}{c} n_2 \end{bmatrix} \quad (5.1.16)$$

With these observations, we can easily show that an alternative definition of  $\mathbf{j}$  may be obtained by noting that  $\mathbf{E}(x_3) = \mathbf{J}(0, x_3)\mathbf{E}(0)$ , where  $x_3 = 0$  denotes the origin of the optical medium. By differentiation, we obtain  $d\mathbf{E}(x_3)/dx_3 = d\mathbf{J}(0, x_3)/dx_3 \mathbf{E}(0)$ . Substitution of Eq. (5.1.14) gives

$$\mathbf{j}(x_3) \equiv \frac{1}{\mathbf{J}(0, x_3)} \frac{d\mathbf{J}(0, x_3)}{dx_3}$$

The optical properties of the anisotropic medium are summarized in the  $\mathbf{j}$ -matrix

$$\frac{d\mathbf{E}(x_3)}{dx_3} = \mathbf{j}(x_3)\mathbf{E}(x_3) \quad (5.1.17)$$

This first-order differential equation governs the evolution of the Jones vector along the direction of propagation. Expanding Eq. (5.1.17), we find that

$$\frac{dE_1}{dx_3} = j_{11}E_1 + j_{12}E_2 \quad (5.1.18a)$$

$$\frac{dE_2}{dx_3} = j_{21}E_1 + j_{22}E_2 \quad (5.1.18b)$$

Equation (5.1.17) can be formally solved using the method of ordered exponential operators. The exponentiated solution is written at depth  $x_3$  as

$$\mathbf{E}(x_3) = \exp\left(\int_0^{x_3} \mathbf{j}(x_3) dx'_3\right) \mathbf{E}(0) \quad (5.1.19)$$

where  $\mathbf{E}(0)$  represents the Jones vector of the incident optical field at  $x_3 = 0$ . Thus the wave evolution is described by a continuous and single-valued function of the initial condition.

Clark Jones was the first to prove that, at most, eight matrices describe all possible types of elementary behaviors that may coexist in an anisotropic dielectric medium [8]. By "simultaneously" we mean that the optical medium has both birefringence and dichroism, in contrast to a cascade of optical elements alternately exhibiting these two properties. This observation is already apparent in the Jones matrix formulation, which contains, at most, four complex elements, that is, eight real independent parameters. The matrix  $\mathbf{j}$  in Eq. (5.1.14) has four complex elements and thus can be expressed as a linear combination in terms of the Pauli matrices  $\sigma_k$ ,

$$\mathbf{j}(x_3) = \sum_{k=0}^3 (a_k(x_3) + b_k(x_3)) \sigma_k \quad (5.1.20)$$

where the  $a_k$  and  $b_k$  coefficients of the expansion are functions of  $x_3$  and depend on birefringence (retardation) and dichroism (attenuation) in the medium. More specifically, the  $a_0$  coefficient represents the arbitrary phase of the optical medium (polarization-independent characteristics) whereas the  $a_1$  coefficient is the linear birefringence in the  $x_1$  direction over that in the  $x_2$  direction, the  $a_2$  coefficient is the linear birefringence at  $45^\circ$  with respect to the  $x_1$  direction over that at  $-45^\circ$  with respect to the  $x_1$  direction and the  $a_3$  coefficient is the right circular birefringence over left circular birefringence.<sup>3</sup> In like fashion, the  $b_0$  coefficient represents the absorption of the optical medium (polarization-independent characteristics) whereas the  $b_1$  coefficient is the linear dichroism in the  $x_1$  direction over that in the  $x_2$  direction, the  $b_2$  coefficient is the linear dichroism at  $45^\circ$  with respect to the  $x_1$  direction over that at  $-45^\circ$  with respect to the  $x_1$  direction, and the  $b_3$  coefficient is the right circular dichroism over left circular dichroism.

Another remarkable result, first discussed by Clark Jones, arises from the relationship between the  $\mathbf{j}$  and  $\mathbf{J}$  matrices [8]. Let us amplify this result. Consider an arbitrary Jones differential polarization matrix

$$\mathbf{j} = \begin{bmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{bmatrix}$$

<sup>3</sup>Circular birefringence induces a differential retardation in the phase of the orthogonal states (right and left) of circularly polarized light. Circular dichroism represents an anisotropic attenuation of left and right circularly polarized light.

By Maclaurin series expansion, Clark Jones was able to prove that

$$\mathbf{J}(0, x_3) = \exp(Cx_3) \begin{bmatrix} \cosh(Ax_3) + \frac{B \sinh(Ax_3)}{A} & j_{12} \frac{\sinh(Ax_3)}{A} \\ j_{21} \frac{\sinh(Ax_3)}{A} & \cosh(Ax_3) - \frac{B \sinh(Ax_3)}{A} \end{bmatrix} \quad (5.1.21)$$

where

$$A \equiv (C^2 - D)^{1/2}, \quad B \equiv \frac{1}{2}(j_{11} - j_{22}), \quad C \equiv \frac{1}{2}(j_{11} + j_{22}) \quad \text{and} \quad D \equiv j_{11}j_{22} - j_{12}j_{21}$$

This equation forms a basis for the formulation of the propagation of waves in anisotropic media.

The calculations we have carried out are inapplicable when the optical medium depolarizes an incident pure state of polarization. In that case a natural formulation of this problem is to assume that the input and output Stokes parameters are linearly related via a Mueller matrix. I assume again that the medium is transversally homogeneous (along the  $x_1$  and  $x_2$  axes) but can be longitudinally inhomogeneous (along the  $x_3$  axis). The analog of the Mueller representation in terms of a differential equation analysis has been discussed by Azzam [3, 9] and may be expressed in the form

$$\mathbf{m}(x_3) \equiv \lim_{\Delta x_3 \rightarrow 0} \frac{\mathbf{M}(x_3, \Delta x_3) - \mathbf{O}_0^{(4)}}{\Delta x_3} \quad (5.1.22)$$

where the optical properties of the anisotropic medium are contained in the  $4 \times 4$  differential Mueller matrix  $\mathbf{m}$ . I leave to the reader the task of verifying that another expression of  $\mathbf{m}$  evaluated at a distance  $x_3$  from the origin of coordinates satisfies

$$\mathbf{m}(x_3) \equiv \frac{1}{\mathbf{M}(0, x_3)} \frac{d\mathbf{M}(0, x_3)}{dx_3} \quad (5.1.23)$$

We observe that Eq. (5.1.23) relates the differential matrix  $\mathbf{m}$  describing the optical properties of a thin slab of the medium to the Mueller matrix  $\mathbf{M}$  describing the optical properties of a thick slab of the same medium. The first-order differential equation governing the evolution of the Stokes vector can be proved in a similar manner as Eq. (5.1.17), and we have

$$\frac{d\mathbf{S}(x_3)}{dx_3} = \mathbf{m}\mathbf{S}(x_3) \quad (5.1.24)$$

where  $\mathbf{S}(x_3)$  is the Stokes vector at  $x_3$ . Notice that Eq. (5.1.25) is equivalent to

four simultaneous first-order linear differential equations in the four instantaneous Stokes parameters.

As an illustrative example, we consider the Mueller matrix of a homogenous linearly birefringent and nonabsorbing plate with fast and slow axes along the  $x_1$  and  $x_2$  axes of the reference coordinate system [i.e., Eq. (4.1.44)]

$$\mathbf{M}(0, x_3) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\left(\frac{2\pi v \Delta n}{c} x_3\right) & \sin\left(\frac{2\pi v \Delta n}{c} x_3\right) \\ 0 & 0 & -\sin\left(\frac{2\pi v \Delta n}{c} x_3\right) & \cos\left(\frac{2\pi v \Delta n}{c} x_3\right) \end{bmatrix} \quad (5.1.25)$$

where  $\Delta n$  denotes the linear birefringence of the medium. By expanding the cosine and sine functions to first order in their power series, we have

$$\mathbf{M}(0, x_3) = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & \frac{2\pi v \Delta n}{c} x_3 \\ 0 & 0 & -\frac{2\pi v \Delta n}{c} x_3 & 1 \end{bmatrix} \quad (5.1.26)$$

It follows that the  $\mathbf{m}$  matrix reads as

$$\mathbf{m} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{2\pi v \Delta n}{c} \\ 0 & 0 & -\frac{2\pi v \Delta n}{c} & 0 \end{bmatrix} \quad (5.1.27)$$

Notice that Eq. (5.1.27) could also have been obtained by substituting Eq. (5.1.25) into Eq. (5.1.23). Before going further, we would like to emphasize that if  $\mathbf{M}$  can be derived from a Jones matrix  $\mathbf{J}$ , namely, what we previously called a *Mueller-Jones matrix*, we can write

$$\mathbf{M}_J(x_3, \Delta x_3) = \mathbf{A}(\mathbf{J}(x_3, \Delta x_3) \otimes \mathbf{J}^*(x_3, \Delta x_3))\mathbf{A}^{-1} \quad (5.1.28)$$

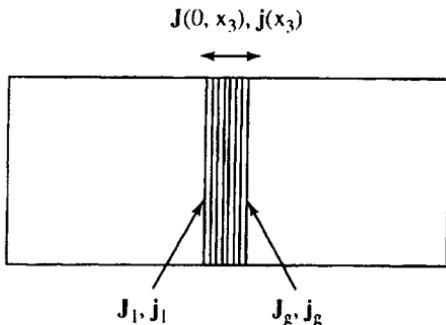
where the constant matrix  $\mathbf{A}$  is as given by Eq. (3.1.171). We can rewrite Eq. (5.1.14) as  $\mathbf{J}(x_3, \Delta x_3) = \sigma_0 + \mathbf{j}(x_3)\Delta x_3$  and substitute into Eq. (5.1.29). This yields a Mueller matrix  $\mathbf{M}_j(x_3, \Delta x_3)$  that can be put in the form  $\mathbf{M}_j(x_3, \Delta x_3) = \mathbf{O}_0^{(4)} + \mathbf{m}_j(x_3)\Delta x_3$  to first order in  $\Delta x_3$ . The differential polarization matrix  $\mathbf{m}_j(x_3)$  can thus be written

$$\mathbf{m}_j(x_3) = \mathbf{A}(\sigma_0 \otimes \mathbf{j}^*(x_3) + \mathbf{j}(x_3) \otimes \sigma_0)\mathbf{A}^{-1} \quad (5.1.29)$$

According to Eq. (5.1.29), we verify that an arbitrary anisotropic Mueller-Jones medium has seven independent real parameters at most. The reader can easily verify that the  $\mathbf{m}$  matrices of the eight types of optical behavior listed in Table 5.1.2 do satisfy Eq. (5.1.29). By substituting Eq. (5.1.20) into Eq. (5.1.29), one may prove that the most general  $\mathbf{m}$  matrix of a Mueller-Jones medium can be obtained by adding the  $\mathbf{m}_k$  matrices that correspond to the different optical properties listed in Table 5.1.2

$$\mathbf{m}_j = 2 \begin{bmatrix} b_0 & b_1 & b_2 & b_3 \\ b_1 & b_0 & -ia_3 & ia_2 \\ b_2 & ia_3 & b_0 & -ia_1 \\ b_3 & -ia_2 & ia_1 & b_0 \end{bmatrix} \quad (5.1.30)$$

Now, if we limit our interest entirely to nondepolarizing optical media, it is useful to devote several paragraphs to a discussion of the lamellae formalism due to Clark Jones [8]. Assume that an infinitesimal layer of the macroscopic sample is made up of a stack of eight lamellae, all of equal thickness (one-eighth that of the layer) and each possessing one and only one of the eight basic properties listed in Table 5.1.2 (see Fig. 5.1.2). In Table 5.1.2, we list the eight basic differential polarization Mueller matrices that enter the lamellae theory of Clark Jones.



**Figure 5.1.2.** Illustrating the lamellae formalism of Clark Jones to represent a material containing a combination of optic anisotropies. Each lamella represents a specific optical effect.

**TABLE 5.1.2.** Differential Polarization Matrices  $j_k$  and  $m_k$  Corresponding to the Eight Elementary Optical Properties into Which an Arbitrary Optical Behavior Can Be Resolved

Optical Property	$m_k$	$j_k$
Isotropic refraction (arbitrary phase)	$m_1 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$j_1 = a_0 \sigma_0$
Isotropic absorption	$m_2 = \begin{bmatrix} 2b_0 & 0 & 0 & 0 \\ 0 & 2b_0 & 0 & 0 \\ 0 & 0 & 2b_0 & 0 \\ 0 & 0 & 0 & 2b_0 \end{bmatrix}$	$j_2 = b_0 \sigma_0$
Linear birefringence along $x_1$ and $x_2$ axes	$m_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2ia_1 \\ 0 & 0 & 2ia_1 & 0 \end{bmatrix}$	$j_3 = a_1 \sigma_1$
Linear dichroism along $x_1$ and $x_2$ axes	$m_4 = \begin{bmatrix} 0 & 2b_1 & 0 & 0 \\ 2b_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$j_4 = b_1 \sigma_1$
Linear birefringence along bisectors of $x_1$ and $x_2$ axes	$m_5 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2ia_2 \\ 0 & 0 & 0 & 0 \\ 0 & -2ia_2 & 0 & 0 \end{bmatrix}$	$j_5 = a_2 \sigma_2$
Linear dichroism along bisectors of $x_1$ and $x_2$ axes	$m_6 = \begin{bmatrix} 0 & 0 & 2b_2 & 0 \\ 0 & 0 & 0 & 0 \\ 2b_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$j_6 = b_2 \sigma_2$
Circular birefringence	$m_7 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -2ia_3 & 0 \\ 0 & 2ia_3 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$j_7 = a_3 \sigma_3$
Circular dichroism	$m_8 = \begin{bmatrix} 0 & 0 & 0 & 2b_3 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2b_3 & 0 & 0 & 0 \end{bmatrix}$	$j_8 = b_3 \sigma_3$

<sup>a</sup>The eight  $j_k$  matrices correspond to the eight  $\theta_k$  matrices given by Clark Jones in his original derivation [8]. Note that the  $m_k$  matrices are Mueller-Jones differential matrices. The eight coefficients are  $a_0 = -i(2\pi v n/c)$ ,  $b_0 = -(2\pi v k/c)$ ,  $a_1 = i(\pi v \Delta n/c)$ ,  $b_1 = -(\pi v \Delta k/c)$ ,  $a_2 = i(\pi v \Delta n_{45}/c)$ ,  $b_2 = -(\pi v \Delta k_{45}/c)$ ,  $a_3 = i(\pi v \Delta k_c/c)$ ,  $b_3 = (\pi v \Delta k_c/c)$ , where  $n$  (resp.  $k$ ) is the real (resp. imaginary) part of the refractive index;  $\Delta n$  (resp.  $\Delta k$ ) is the linear birefringence, which is along the  $x_1$  and  $x_2$  axes (resp. the linear dichroism, which is along the  $x_1$  and  $x_2$  axes);  $\Delta n_{45}$  (resp.  $\Delta k_{45}$ ) is the linear birefringence, which is along the bisectors of the  $x_1$  and  $x_2$  axes (resp. the linear dichroism, which is along the bisectors of the  $x_1$  and  $x_2$  axes);  $\Delta n_c$  (resp.  $\Delta k_c$ ) is the circular birefringence (resp. the circular dichroism).

On expressing the Jones matrix  $\mathbf{J}$  for this planar stratified medium in terms of the Jones matrices of the individual optical elements  $\mathbf{J}_k$  via Eq. (4.1.70) and making use of Eq. (5.1.19), which can be rewritten as  $\mathbf{J}(0, x_3) = \exp(\int_0^{x_3} \mathbf{j}(x'_3) dx'_3)$ , we can write the differential Jones matrix  $\mathbf{j}$  as a linear superposition of the eight matrices  $\mathbf{j}_k$ :

$$\mathbf{j}(x_3) = \sum_{k=1}^8 \mathbf{j}_k \quad (5.1.31)$$

The differential matrices for lamellae, each characterizing a specific optical effect, are simply additive. Whereas Jones matrices are noncommuting exponential operators, it is interesting to observe that the ordering of the lamellae does not affect the final result as required.

To illustrate these results, consider the special but important case of a material with coaxial linear birefringence and linear dichroism. From Table 5.1.2, we can write  $\mathbf{j} = (a_1 + b_1)\boldsymbol{\sigma}_1$ . Consequently, the Jones matrix follows directly from Eq. (5.1.21) and is

$$\mathbf{J}(0, x_3) = \begin{bmatrix} \exp\left(\frac{i\pi\nu x_3}{c}(\Delta n + i\Delta k)\right) & 0 \\ 0 & \exp\left(-\frac{i\pi\nu x_3}{c}(\Delta n + i\Delta k)\right) \end{bmatrix} \quad (5.1.32)$$

In this section we assumed that the wave propagates in a nonmagnetic material having anisotropic permittivity but scalar permeability. We would like to note that magnetic materials are usually anisotropic as a consequence of the microscopic mechanisms that underlie magnetism. The case in which the anisotropic material has a scalar permittivity but a tensor permeability has been considered by Vernon and Huggins [11].

### 5.1.3. ELLIPSE OF POLARIZATION IN ANISOTROPIC MEDIA

The procedure we have just outlined, leading to the first-order differential equation for the Jones vector [i.e., Eq. (5.1.17)], can be complemented by the complex polarization ratio representation of polarized light. Information on the polarization ellipse can be extracted from the Jones vector by taking the ratio of its two components. See Section 3.1.4 for details. The first-order differential equation governing the evolution of the polarization ellipse in homogeneous anisotropic media whose optical properties depend only on  $x_3$  was first derived by Azzam and Bashara [3, 10]. We do not duplicate the analysis here, but several results are worth stating. Let us consider a plane wave propagating along the  $x_3$ -axis of the  $(x_1, x_2, x_3)$  Cartesian coordinate system. The derivation of the first-order nonlinear differential equation that governs the evolution of the complex number  $Z_{12} = E_2/E_1$  as the lightwave propagates through the anisotropic medium is as follows. If first differentiate

Eq. (3.1.50) with respect to  $x_3$ . The result is

$$\frac{dZ_{12}(x_3)}{dx_3} = \frac{1}{E_1} \frac{dE_2}{dx_3} - \frac{E_2}{E_1^2} \frac{dE_1}{dx_3} \quad (5.1.33)$$

On substitution from Eqs. (5.1.18a, b) and (3.1.50) into Eq. (5.1.33), we find that

$$\frac{dZ_{12}(x_3)}{x_3} = -j_{12}(Z_{12}(x_3))^2 + (j_{22} - j_{11})Z_{12}(x_3) + j_{21} \quad (5.1.34)$$

This relation reduces to a usual Riccati equation whose solution can be determined from the initial condition  $Z_{12}(x_3 = 0) = Z_0$ . Consequently it is possible to find  $Z_{12}(x_3)$ , for a given direction of propagation, when the matrix  $\mathbf{j}$  takes on any arbitrary form. As an exercise, the reader can readily show that for a homogeneous anisotropic crystal, that is, where every element of the matrix  $\mathbf{j}$  is constant independent of  $x_3$ , the corresponding solution of Eq. (5.1.34) reads as

$$Z_{12}(x_3) = \frac{aZ_0 + \left( \frac{Z_0(j_{22} - j_{11})}{2} + j_{21} \right) \tan(ax_3)}{a + \left( j_{12}Z_0 + \frac{(j_{11} - j_{22})}{2} \right) \tan(ax_3)} \quad (5.1.35)$$

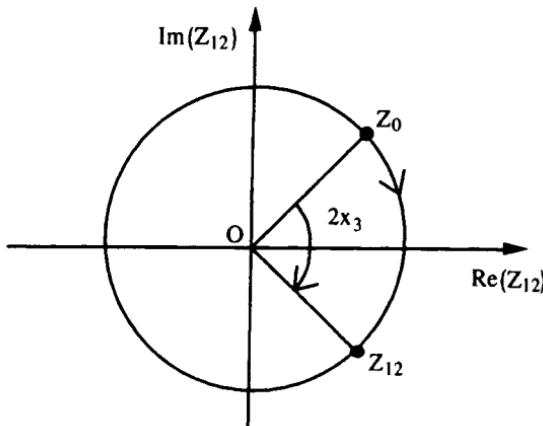
where we have set  $a = (-\frac{1}{4}(j_{11} - j_{22})^2 - j_{12}j_{21})^{1/2}$ . In the general case, it is useful to represent this solution by a trajectory in the complex plane as the distance of propagation,  $x_3$ , is increased from an initial state of polarization  $Z_0$  at  $x_3 = 0$ . By virtue of the stereographic projection, this trajectory can be also visualized as a trajectory on the Poincaré sphere  $\Sigma_1^2$ . Bear in mind that  $Z_{12}$  has information on the azimuth  $\psi$  and ellipticity  $\chi$  of the polarization ellipse: the corresponding spatial variations  $\psi(x_3)$  and  $\chi(x_3)$  are given from Eqs. (3.1.51a) and (3.1.51b), respectively.

Since our results to this point have all been stated in fairly general terms, it may be of help to discuss a specific situation. Let us consider, as a particularly simple example, a homogeneous linearly birefringent medium whose principal axes of birefringence are parallel to the  $x_1$  and  $x_2$  directions. For this material, the  $\mathbf{j}$  matrix reads as

$$\mathbf{j} = ig_0 \boldsymbol{\sigma}_1 \quad (5.1.36)$$

where  $g_0 = \pi v \Delta n / c$  is proportional to the difference between the two indices of refraction along the  $x_1$  and  $x_2$  axes. From Eq. (5.1.35) it follows directly that

$$Z_{12} = \exp(-2ig_0 x_3) Z_0 \quad (5.1.37)$$

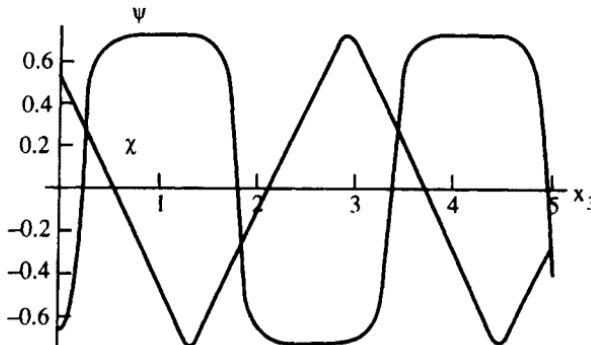


**Figure 5.1.3.** The trajectory for light propagating into a homogeneous linearly birefringent medium of differential polarization matrix  $\mathbf{j} = ig_o\sigma_1$ .

Thus the trajectory of  $Z_{12}(x_3)$  shown in the complex plane is a circle with center at the origin (Fig. 5.1.3) of the plane. As the distance of propagation  $x_3$  is increased, the motion of the representative point  $Z_{12}$  is a periodic function of  $x_3$  with period  $\pi/g_o$ .

In accordance with our previous remark, the variation of  $\psi(x_3)$  and of  $\chi(x_3)$  can be obtained by substituting Eq. (5.1.37) into Eqs. (3.1.51a, b), leaving

$$\tan(2\psi) = \frac{2(\cos(2g_o x_3)\operatorname{Re}(Z_0) - \sin(2g_o x_3)\operatorname{Im}(Z_0))}{1 - |Z_0|^2} \quad (5.1.38a)$$



**Figure 5.1.4.** Plot showing the evolution of the azimuth  $\psi$  and the ellipticity  $\chi$ , as a function of distance  $x_3$ , of a pure state of polarization propagating into a homogeneous birefringent medium of differential polarization matrix  $\mathbf{j} = ig_o\sigma_1$  with  $Z_0 = 0.5 + i$  and  $0 \leq g_o x_3 \leq 5$ .

and

$$\sin(2\chi) = \frac{2(\cos(2g_0x_3)\text{Im}(Z_0) + \sin(2g_0x_3)\text{Re}(Z_0))}{1 + |Z_0|^2} \quad (5.1.38b)$$

A plot of the azimuth and ellipticity with distance  $x_3$  is shown in Fig. 5.1.4. The example just described concerned a homogeneous anisotropic medium. For an inhomogeneous anisotropic medium, the elements of the  $\mathbf{j}$  matrix are functions of  $x_3$ , the distance along the direction of propagation. In the general case the differential equation (5.1.33) does not have an analytic solution [11].

#### 5.1.4. EVOLUTION OF STOKES PARAMETERS IN ANISOTROPIC MEDIA

The purpose of this section is to derive nonlinear differential equations that express the spatial variations of the normalized Stokes parameters as the light propagates into an optically anisotropic medium. With the help of the Stokes vector space, we can then visually grasp the continuous complicate change in the state of polarization along a given direction of propagation.

Start from Eqs. (5.1.18a, b). The first step toward obtaining differential equations in terms of the Stokes parameters consists of multiplying Eq. (5.1.18a) by  $E_1^*$  and the complex conjugate of Eq. (5.1.18a) by  $E_1$ , then taking the sum and time averaging. This gives

$$\frac{d\langle S_0 \rangle}{dx_3} + \frac{d\langle S_1 \rangle}{dx_3} = 2\text{Re}(j_{11})(\langle S_0 \rangle + \langle S_1 \rangle) + 2\text{Re}(j_{12})(\langle S_2 \rangle) - 2\text{Im}(j_{12})(\langle S_3 \rangle) \quad (5.1.39a)$$

In like fashion, it is easily verified that

$$\frac{d\langle S_0 \rangle}{dx_3} - \frac{d\langle S_1 \rangle}{dx_3} = 2\text{Re}(j_{22})(\langle S_0 \rangle - \langle S_1 \rangle) + 2\text{Re}(j_{21})(\langle S_2 \rangle) + 2\text{Im}(j_{21})(\langle S_3 \rangle) \quad (5.1.39b)$$

$$\begin{aligned} \frac{d\langle S_2 \rangle}{dx_3} - i \frac{d\langle S_3 \rangle}{dx_3} &= (j_{11} + j_{22}^*)(\langle S_2 \rangle - i\langle S_3 \rangle) + j_{12}(\langle S_0 \rangle - \langle S_1 \rangle) \\ &\quad + j_{21}^*(\langle S_0 \rangle + \langle S_1 \rangle) \end{aligned} \quad (5.1.39c)$$

$$\begin{aligned} \frac{d\langle S_2 \rangle}{dx_3} + i \frac{d\langle S_3 \rangle}{dx_3} &= j_{21}(\langle S_0 \rangle + \langle S_1 \rangle) + j_{12}^*(\langle S_0 \rangle - \langle S_1 \rangle) \\ &\quad + (j_{11}^* + j_{22})(\langle S_2 \rangle + i\langle S_3 \rangle) \end{aligned} \quad (5.1.39d)$$

From Eqs. (5.1.39a–d), we recognize that

$$\begin{aligned} \frac{1}{\langle S_0 \rangle} \frac{d\langle S_0 \rangle}{dx_3} &= \operatorname{Re}(j_{11})(1 + \langle \sigma_1 \rangle) + \operatorname{Re}(j_{22})(1 - \langle \sigma_1 \rangle) + \operatorname{Re}(j_{12})\langle \sigma_2 \rangle \\ &\quad + \operatorname{Re}(j_{21})\langle \sigma_2 \rangle - \operatorname{Im}(j_{12})\langle \sigma_3 \rangle + \operatorname{Im}(j_{21})\langle \sigma_3 \rangle \end{aligned} \quad (5.1.40a)$$

$$\begin{aligned} \frac{1}{\langle S_0 \rangle} \frac{d\langle S_1 \rangle}{dx_3} &= \operatorname{Re}(j_{11})(1 + \langle \sigma_1 \rangle) - \operatorname{Re}(j_{22})(1 - \langle \sigma_1 \rangle) + \operatorname{Re}(j_{12})\langle \sigma_2 \rangle \\ &\quad - \operatorname{Re}(j_{21})\langle \sigma_2 \rangle - \operatorname{Im}(j_{12})\langle \sigma_3 \rangle - \operatorname{Im}(j_{21})\langle \sigma_3 \rangle \end{aligned} \quad (5.1.40b)$$

$$\begin{aligned} \frac{1}{\langle S_0 \rangle} \frac{d\langle S_2 \rangle}{dx_3} &= \operatorname{Re}(j_{21})(1 + \langle \sigma_1 \rangle) + \operatorname{Re}(j_{12})(1 - \langle \sigma_1 \rangle) + (\operatorname{Re}(j_{11}) + \operatorname{Re}(j_{22}))\langle \sigma_2 \rangle \\ &\quad + (\operatorname{Im}(j_{11})\langle \sigma_3 \rangle - \operatorname{Im}(j_{22}))\langle \sigma_3 \rangle \end{aligned} \quad (5.1.40c)$$

$$\begin{aligned} \frac{1}{\langle S_0 \rangle} \frac{d\langle S_3 \rangle}{dx_3} &= \operatorname{Re}(j_{21})\langle \sigma_3 \rangle + \operatorname{Re}(j_{12})\langle \sigma_3 \rangle + \operatorname{Im}(j_{21})(1 + \langle \sigma_1 \rangle) \\ &\quad + \operatorname{Im}(j_{12})(-1 + \langle \sigma_1 \rangle) + (-\operatorname{Im}(j_{11}) + \operatorname{Im}(j_{22}))\langle \sigma_2 \rangle \end{aligned} \quad (5.1.40d)$$

Now from the very definition of the normalized Stokes parameters  $\langle \sigma_j \rangle \equiv \langle S_j \rangle / \langle S_0 \rangle$ , we have

$$\frac{d\langle \sigma_j \rangle}{dx_3} \equiv \frac{1}{\langle S_0 \rangle} \frac{d\langle S_j \rangle}{dx_3} - \frac{\langle \sigma_j \rangle}{\langle S_0 \rangle} \frac{d\langle S_0 \rangle}{dx_3} \quad (5.1.41)$$

By substituting Eqs. (5.1.40a, b) into Eqs. (5.1.41), we can derive nonlinear differential equations that express the spatial variations of the normalized Stokes parameters as the light propagates into the optically anisotropic medium, starting from the initial state of polarization  $(\langle \sigma_1 \rangle_i, \langle \sigma_2 \rangle_i, \langle \sigma_3 \rangle_i)$  at  $x_3 = 0$ :

$$\begin{aligned} \frac{d\langle \sigma_1 \rangle}{dx_3} &= \operatorname{Re}(j_{11})(1 - \langle \sigma_1 \rangle^2) + \operatorname{Re}(j_{22})(\langle \sigma_1 \rangle^2 - 1) + \operatorname{Re}(j_{12})\langle \sigma_2 \rangle(1 - \langle \sigma_1 \rangle) \\ &\quad - \operatorname{Re}(j_{21})\langle \sigma_2 \rangle(1 + \langle \sigma_1 \rangle) + \operatorname{Im}(j_{12})\langle \sigma_3 \rangle(\langle \sigma_1 \rangle - 1) \\ &\quad - \operatorname{Im}(j_{21})\langle \sigma_3 \rangle(\langle \sigma_1 \rangle + 1) \end{aligned} \quad (5.1.42a)$$

$$\begin{aligned} \frac{d\langle \sigma_2 \rangle}{dx_3} &= -\operatorname{Re}(j_{11})\langle \sigma_1 \rangle\langle \sigma_2 \rangle + \operatorname{Re}(j_{22})\langle \sigma_1 \rangle\langle \sigma_2 \rangle - \operatorname{Im}(j_{22})\langle \sigma_3 \rangle + \operatorname{Im}(j_{11})\langle \sigma_3 \rangle \\ &\quad + \operatorname{Re}(j_{12})(1 - \langle \sigma_1 \rangle - \langle \sigma_2 \rangle^2) + \operatorname{Re}(j_{21})(1 + \langle \sigma_1 \rangle - \langle \sigma_2 \rangle^2) \\ &\quad + \operatorname{Im}(j_{12})\langle \sigma_2 \rangle\langle \sigma_3 \rangle - \operatorname{Im}(j_{21})\langle \sigma_2 \rangle\langle \sigma_3 \rangle \end{aligned} \quad (5.1.42b)$$

$$\begin{aligned} \frac{d\langle\sigma_3\rangle}{dx_3} = & -\operatorname{Re}(j_{11})\langle\sigma_1\rangle\langle\sigma_3\rangle + \operatorname{Re}(j_{22})\langle\sigma_1\rangle\langle\sigma_3\rangle + \operatorname{Im}(j_{22})\langle\sigma_2\rangle - \operatorname{Im}(j_{11})\langle\sigma_2\rangle \\ & -\operatorname{Re}(j_{12})\langle\sigma_2\rangle\langle\sigma_3\rangle - \operatorname{Re}(j_{21})\langle\sigma_2\rangle\langle\sigma_3\rangle + \operatorname{Im}(j_{12})(\langle\sigma_3\rangle^2 - 1 + \langle\sigma_1\rangle) \\ & + \operatorname{Im}(j_{21})(1 + \langle\sigma_1\rangle - \langle\sigma_3\rangle^2) \end{aligned} \quad (5.1.42c)$$

These expressions are the basic differential equations that describe the propagation of the Stokes parameters in an optically anisotropic Jones medium. These nonlinear equations offer a rich store of valuable information depending on the particular choice of the  $\mathbf{j}$  matrix.

In the following, we present the detailed solution to the set of nonlinear equations for an example chosen to illustrate the generality of the preceding analysis. Let us return to the previously considered  $\mathbf{j}$  matrix  $\mathbf{j} = ig_o\sigma_1$ , corresponding to light propagation along the  $x_3$  axis in a homogeneous linearly birefringent medium whose principal axes are parallel to the  $x_1$  and  $x_2$  directions. We easily see that the general solution of Eqs. (5.1.42a–c) is given by

$$\langle\sigma_1\rangle = \langle\sigma_1\rangle_i \quad (5.1.43a)$$

$$\langle\sigma_2\rangle = \langle\sigma_3\rangle_i \sin(2g_o x_3) + \langle\sigma_2\rangle_i \cos(2g_o x_3) \quad (5.1.43b)$$

$$\langle\sigma_3\rangle = -\langle\sigma_2\rangle_i \sin(2g_o x_3) + \langle\sigma_3\rangle_i \cos(2g_o x_3) \quad (5.1.43c)$$

Take, for example a linear horizontally polarized light,  $\langle\sigma_1\rangle_i = 1$ ,  $\langle\sigma_2\rangle_i = 0$ ,  $\langle\sigma_3\rangle_i = 0$ . A sketch of the trajectory, represented as a parametric plot of the  $\langle\sigma_2\rangle$  and  $\langle\sigma_3\rangle$  coordinates, is given in Fig. 5.1.5.

We next consider a time-independent  $\mathbf{m}$  matrix in Eq. (5.1.24). From Eq. (5.1.25), we can derive nonlinear first-order differential equations that express the spatial variations of the normalized Stokes parameters as the light propagates into the optically anisotropic medium, starting from the initial state of polarization ( $\langle\sigma_1\rangle_i$ ,  $\langle\sigma_2\rangle_i$ ,  $\langle\sigma_3\rangle_i$ ) at  $x_3 = 0$ :

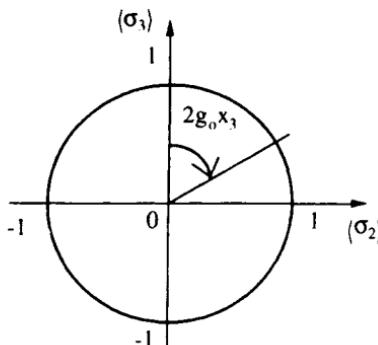
$$\frac{1}{S_0} \frac{dS_0}{dx_3} = n_{00} + n_{01}\langle\sigma_1\rangle + n_{02}\langle\sigma_2\rangle + n_{03}\langle\sigma_3\rangle \quad (5.1.44a)$$

$$\frac{1}{S_0} \frac{dS_1}{dx_3} = n_{10} + n_{11}\langle\sigma_1\rangle + n_{12}\langle\sigma_2\rangle + n_{13}\langle\sigma_3\rangle \quad (5.1.44b)$$

$$\frac{1}{S_0} \frac{dS_2}{dx_3} = n_{20} + n_{21}\langle\sigma_1\rangle + n_{22}\langle\sigma_2\rangle + n_{23}\langle\sigma_3\rangle \quad (5.1.44c)$$

$$\frac{1}{S_0} \frac{dS_3}{dx_3} = n_{30} + n_{31}\langle\sigma_1\rangle + n_{32}\langle\sigma_2\rangle + n_{33}\langle\sigma_3\rangle \quad (5.1.44d)$$

Here the  $n_{kl}$  terms denote the matrix elements of  $\mathbf{m}$ .



**Figure 5.1.5.** The trajectory, represented as a parametric plot of the  $\langle \sigma_2 \rangle$  and  $\langle \sigma_3 \rangle$  coordinates that describes the propagation of the state of polarization in a homogeneous linearly birefringent medium of differential polarization matrix  $\mathbf{j} = ig_0\sigma_1$ .

By substituting Eqs. (5.1.44a–d) into Eq. (5.1.41), we can derive a set of three first-order nonlinear differential equations in the form

$$\begin{aligned} \frac{d\langle \sigma_1 \rangle}{dx_3} &= n_{10} + (n_{11} - n_{00})\langle \sigma_1 \rangle - n_{01}\langle \sigma_1 \rangle^2 + n_{12}\langle \sigma_2 \rangle + n_{13}\langle \sigma_3 \rangle \\ &\quad - n_{02}\langle \sigma_1 \rangle\langle \sigma_2 \rangle - n_{03}\langle \sigma_1 \rangle\langle \sigma_3 \rangle \end{aligned} \quad (5.1.45a)$$

$$\begin{aligned} \frac{d\langle \sigma_2 \rangle}{dx_3} &= n_{20} + (n_{22} - n_{00})\langle \sigma_2 \rangle - n_{02}\langle \sigma_2 \rangle^2 + n_{21}\langle \sigma_1 \rangle + n_{23}\langle \sigma_3 \rangle \\ &\quad - n_{01}\langle \sigma_1 \rangle\langle \sigma_2 \rangle - n_{03}\langle \sigma_2 \rangle\langle \sigma_3 \rangle \end{aligned} \quad (5.1.45b)$$

$$\begin{aligned} \frac{d\langle \sigma_3 \rangle}{dx_3} &= n_{30} + (n_{33} - n_{00})\langle \sigma_3 \rangle - n_{03}\langle \sigma_3 \rangle^2 + n_{31}\langle \sigma_1 \rangle + n_{32}\langle \sigma_2 \rangle \\ &\quad - n_{01}\langle \sigma_1 \rangle\langle \sigma_3 \rangle - n_{02}\langle \sigma_2 \rangle\langle \sigma_3 \rangle \end{aligned} \quad (5.1.45c)$$

Suppose that we wish to compute the trajectory as a parametric plot in the  $(\langle \sigma_1 \rangle, \langle \sigma_2 \rangle, \langle \sigma_3 \rangle)$  coordinates that describes the propagation of the state of polarization in an optically anisotropic medium. From the preceding, we see that we should first specify the differential Mueller matrix  $\mathbf{m}$ , and then integrate Eqs. (5.1.45a–c) over a specified range of  $x_3$ . In addition, we should note that this approach permits us to evaluate the depolarization induced by propagation of the wavefield. To this end, the spatial derivative of  $P$  can be written as

$$\frac{dP}{dx_3} = \frac{1}{P} \left( \langle \sigma_1 \rangle \frac{d\langle \sigma_1 \rangle}{dx_3} + \langle \sigma_2 \rangle \frac{d\langle \sigma_2 \rangle}{dx_3} + \langle \sigma_3 \rangle \frac{d\langle \sigma_3 \rangle}{dx_3} \right) \quad (5.1.46)$$

By substituting Eqs. (5.1.45a–d) into Eq. (5.1.46), we obtain the spatial

variation of the degree of polarization of the light propagating in the anisotropic medium. It is also worth remarking that the different couplings  $\langle \sigma_i \rangle \langle \sigma_k \rangle$  in Eqs. (5.1.45a–d) allow energy transfer between polarization states.

Thus far we have considered optical media that can be described by deterministic Mueller transformation matrices. The differential equations that describe the evolution of the Stokes parameters can be recast following the analysis developed in Section 4.4.1, to cope with the effects of fluctuations in a linear optical medium on a partially polarized wave. In this case we can still describe the evolution of the Stokes vector by Eq. (5.1.24) but the deterministic matrix  $\mathbf{m}$  is changed into

$$\bar{\mathbf{m}} = \mathbf{m} + \langle \mathbf{m}_R(t) \rangle \quad (5.1.47)$$

where we suppose that the sample realization  $\mathbf{m}_R(t)$  is a stationary ergodic process [12]. We make the additional assumption that the fluctuations in the medium are much slower than the fluctuations of the wave, so that these fluctuations are uncorrelated. A practical use of this formalism for describing the evolution of polarization in liquid crystals is now given.

### 5.1.5. APPLICATION TO LIQUID CRYSTALS

Since their initial discovery in 1888 by the Austrian botanist Friedrich Reinitzer, liquid crystalline materials have been investigated intensely for their scientific and technological potentials. For instance, de Gennes and Prost published an authoritative theoretical treatise of the many advances in the understanding of the structure and properties of these materials, and we refer the interested reader to this book for additional details [13]. Because liquid crystals (LC) play an especially important role in display technology, we now devote a short discussion to some of their properties. LC are anisotropic fluids that behave like a liquid and also show many aspects of orientational long-range order that are characteristic of a solid crystal. The result is a material in which elongated molecules maintain orientational correlations over distances many times the molecular dimensions. These molecules are usually elongated, and many LC physical characteristics are anisotropic. Liquid crystals form various phases with various orientations. Of these mesophases, the best known is the nematic phase, in which the rod-shaped molecules are translationally disordered but align on average along a common direction called the director.<sup>4</sup> At least three properties make them ideally suited for

<sup>4</sup>The term *nematic* is derived from the Greek word for thread, and refers to the threadlike defects observed in the director when a sample is viewed in a microscope with crossed polarizers. The term *smectic* is derived from the Greek word for soap, as smectic phases have material properties similar to those of soap, as smectic phases have material properties similar to those of soap. The term *cholesteric* was chosen because various materials closely related to cholesterol can take on this phase.

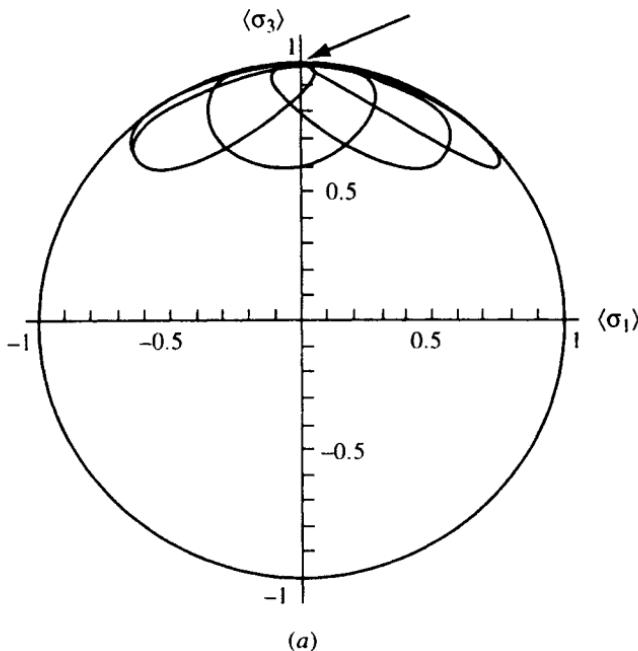
display applications: birefringence, elastic response to orientational deformations, and tendency to align parallel to an applied electric field. Cholesteric mesophases have molecules that are helically organized, and smectic mesophases arrange themselves into distinctive layers, although there is little positional ordering within the layers.<sup>4</sup> Since the intermolecule bond energy is low, the molecules can be easily influenced by external effects such as electric fields. By applying electric or magnetic fields, we can change the orientation of these molecules and consequently cause dramatic changes (phase transitions) in the optical properties of the LC material. Another set of observations involves the use of polarized light. We have seen that crystals can alter polarization of light, and the effect changes dramatically when a crystal sample is rotated with respect to the light direction of propagation. On the other hand, most liquids have no effect on polarization. Thus it was an important discovery that LC could produce changes in light polarization. Polarized light microscopy observations constituted the first clear evidence that such materials are not isotropic but contain some kind of structural order. Electrooptical properties of such materials are important for information display systems, such as calculators and portable computers. A common display method has been the twisted nematic, in which a nematic phase LC is confined between glass sheets treated to align the director with a  $\pi/2$  radian twist between one sheet and the other. Crossed polarizers are placed above and below, and a reflector beneath. When no electric field is applied, the twisted nematic rotates the polarization of light so that the light is reflected, giving the display a silvery appearance. When an electric field is applied, the alignment of the nematic changes the polarizing effect so that light is absorbed rather than reflected, and the display appears to be black [14].

Many of the ideas mentioned above can be illustrated with a simple example, adapted from Ref. [10]. To this end, we examine the detailed solution to Eqs. (5.1.45a–c) for the propagation of a pure state of polarization along the helical axis (assumed here to be  $x_3$ ) of a cholesteric LC. The cholesteric phase displays local alignment of molecules as in the nematic, but with the local director orientation undergoing a periodic twist along one direction. The underlying physics necessary to understand this material can be found in Refs. [13–15]. This system was first treated theoretically by Azzam and Bashara, albeit by using the complex polarization ratio representation of polarized light (see Section 5.1.3) [10].

We assume that the differential Mueller matrix of such an inhomogeneous anisotropic medium takes the following form in the  $(x_1, x_2, x_3)$  Cartesian coordinate system

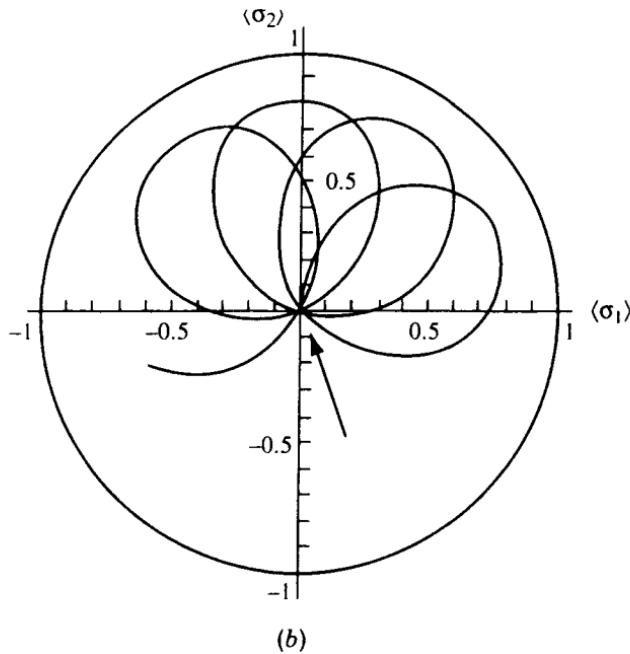
$$\mathbf{m} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g \sin(2\alpha x_3) \\ 0 & 0 & 0 & g \cos(2\alpha x_3) \\ 0 & -g \sin(2\alpha x_3) & -g \cos(2\alpha x_3) & 0 \end{bmatrix} \quad (5.1.48)$$

where  $g$  is a measure of the birefringence and  $\alpha \equiv 2\pi/d$ , in which  $d$  is the pitch of the helical structure of the liquid crystal, that is, the smallest distance between two molecular planes of the same orientation. This result can be understood physically when we realize that each molecular plane acts as a very thin linearly birefringent plate [10]. The principal axes of birefringence, which are determined by the direction of molecular orientation, are gradually rotated from one plane to the next as we proceed along the helical axis of the cholesteric structure. The principal axes of the first molecular plane are aligned with the  $x_1$  and  $x_2$  axes of the coordinate system. The  $\mathbf{m}$  matrix at a distance  $x_3$  is  $\mathbf{m} = \mathbf{R}(-\alpha x_3) \mathbf{m}_3 \mathbf{R}(\alpha x_3)$ , where  $\mathbf{R}$  is the rotation matrix, Eq. (4.1.116) and  $\mathbf{m}_3$  is the matrix that represents linear birefringence along the coordinate axes



(a)

**Figure 5.1.6.** (a) The trajectory, represented as a parametric plot of the  $\langle\sigma_1\rangle$  and  $\langle\sigma_3\rangle$  coordinates, that describe the propagation of the state of polarization in a uniform-pitch cholesteric liquid crystal [Eq. (5.1.48)], with  $g = 1$  and  $0 < \alpha x_3 < 12$ , starting from an incident circular polarization state ( $\langle\sigma_1\rangle_i = 0$ ,  $\langle\sigma_2\rangle_i = 0$ ,  $\langle\sigma_3\rangle_i = 1$ ) indicated by the arrow. The circle  $\langle\sigma_1\rangle^2 + \langle\sigma_3\rangle^2 = 1$  is the circular boundary separating the physical (interior) and the nonphysical (exterior) states of polarization. (b) The trajectory, represented as a parametric plot of the  $\langle\sigma_1\rangle$  and  $\langle\sigma_2\rangle$  coordinates, that describes the propagation of the state of polarization in a uniform-pitch cholesteric liquid crystal [(Eq. (5.1.48))], with  $g = 1$  and  $0 < \alpha x_3 < 12$ , starting from an incident circular polarization state ( $\langle\sigma_1\rangle_i = 0$ ,  $\langle\sigma_2\rangle_i = 0$ ,  $\langle\sigma_3\rangle_i = 1$ ) indicated by the arrow. The circle  $\langle\sigma_1\rangle^2 + \langle\sigma_2\rangle^2 = 1$  is the circular boundary separating the physical (interior) and the nonphysical (exterior) states of polarization.



**Figure 5.1.6.** (Continued).

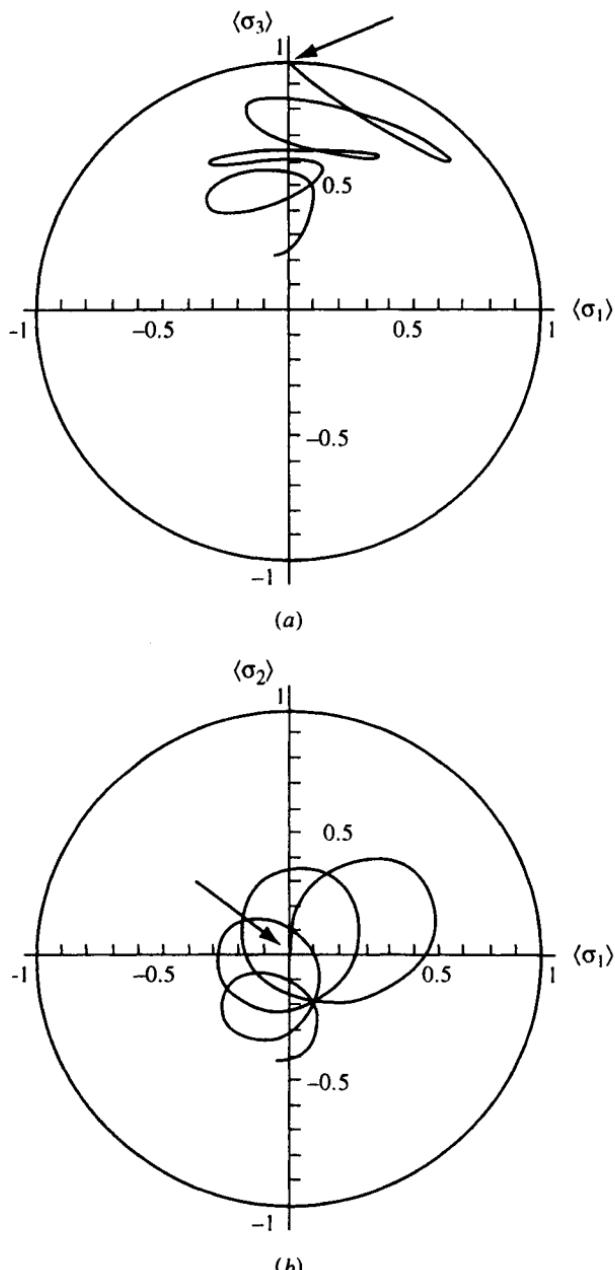
listed in Table 5.1.2. Note that the four nonzero elements of the  $\mathbf{m}$  matrix that appear in Eq. (5.1.48) are functions of  $x_3$ .

The differential equations (5.1.45a–c) were numerically integrated with a fourth-order Runge–Kutta integration scheme [15]. In Fig. 5.1.6a we give the  $\langle\sigma_1\rangle$  and  $\langle\sigma_3\rangle$  coordinates of each point as a function of the parameter  $x_3$ , for an incident circularly polarized beam. Figure 5.1.6b shows the corresponding result for  $\langle\sigma_1\rangle$  and  $\langle\sigma_2\rangle$ . The petal plots in Figs. 5.1.6a,b describe the trajectory of the polarization state as the beam propagates through the medium. The symmetry of these plots reflects the molecular ordering of the cholesteric liquid crystal.

As an example in studying the influence of temporal randomness in the optical medium on the evolution of Stokes parameters, the following Mueller matrix has been chosen:

$$\langle \mathbf{m}_R(t) \rangle = \begin{bmatrix} a & 0 & 0 & 0 \\ 0 & b & 0 & e \\ 0 & 0 & c & e \\ 0 & e & e & d \end{bmatrix} \quad (5.1.49)$$

where the elements  $a$ ,  $b$ ,  $c$ , and  $d$  are real numbers. Physically such additive anisotropic randomness can describe the perturbation in the molecular



**Figure 5.1.7.** (a) Same as Fig. 5.1.6a but with a random perturbation [Eq. (5.1.49)] added to Eq. (5.1.48). The matrix elements in Eq. (5.1.48) are  $a = 0$ ,  $b = -0.04$ ,  $c = -0.09$ ,  $d = 0.09$ , and  $e = 0.02$ . (b) Same as Fig. 5.1.6b but with a random perturbation [Eq. (5.1.49)] added to Eq. (5.1.48). The matrix elements in Eq. (5.1.48) are  $a = 0$ ,  $b = 0.05$ ,  $c = -0.09$ ,  $d = 0.07$ , and  $e = 0.09$ .

ordering of the liquid crystal due to topological defects, such as disinclination loops and boojums [13]. In the presence of the additive randomness, wave energy initially localized in well-defined trajectories in the Poincaré space is significantly delocalized via the mechanism of energy couplings. Figure 5.1.7a shows how the symmetry of these trajectories is broken.

Further, we point out for emphasis that we have only examined the direct problem, that is, to predict the trajectory represented as a parametric plot in the  $(\langle \sigma_1 \rangle, \langle \sigma_2 \rangle, \langle \sigma_3 \rangle)$  coordinates, which describes the propagation of the state of polarization in an optically anisotropic medium, on the basis of a known differential polarization Mueller matrix. No attempt will be made to discuss the inverse problem, specifically, to deduce features of the temporal randomness from the trajectories of Stokes parameters in Poincaré space. A primary goal for the experimentalist is the extraction of information from laboratory data. There are many methods to find solutions to this inverse problem, although it is known that inversion of virtually any type of data has proved to be a notoriously difficult task, in terms of uniqueness and stability. The interested reader should consult Ref. [15].

In a similar vein, the physics that has been presented in the preceding sections has clear relevance to the analysis of the polarization properties of optical fibers. Knowledge of the polarization evolution in optical fibers is important in many applications such as precision sensors, high-speed optical communications, and computations [16]. The maintenance of polarization in anisotropic single-mode optical fibers is essential for coherent optical transmission systems. Polarization degradation may arise from different causes, such as dielectric imperfections of the fiber, scattering, external mechanical stresses, or change in ambient temperature and can introduce a large number of errors in the transmitted information. It is by now well established that variations of anisotropic parameters in optical fibers leads to modulation of the polarization characteristics of the output light. Several reviews in the area of guided-wave optics have now been published and the reader is referred to these for historical background and bibliography [17–21].

The variety of techniques brought to bear on the class of problems involving anisotropic media is a rich one. This chapter has provided a general introduction to study of the propagation of a partially polarized wave in these media. In so doing, I hope to have given to the reader a taste of the conceptual issues that optical physicists are grappling with.

## SECTION 5.2

# Optical Polarizing Components

Since the initial Bartholinus's report of double refraction, crystallographers and geologists have extensively examined the optical properties of crystals by studying their polarization characteristics. In earlier days, the knowledge of these systems was mainly empirical. Now, a much more fundamental and comprehensive picture has emerged, to the point where general principles, connecting the crystal structure to the properties of light polarization, can be used to understand and control optical devices. Ramachandran and Rama-seshan provide an extensive review of crystal optics [6].

In Part 4, the Jones and Mueller matrices were derived for common optical polarizing components, such as polarizers, retarders, and rotators, that modify the state of polarization of light. The basic objective of this chapter is to present a discussion of basic optical polarizing components for the control and detection of the state of polarization of light. Optical polarizing components are elements that modify the state of polarization of light. High-quality components are crucial for high-precision polarimetry. Generally these components utilize the properties of natural or induced birefringence, and dichroism and are made of anisotropic materials. As we have seen, this produces different velocities for orthogonal polarizations depending on the propagation direction in the crystal. *Birefringence* is the material property that results in a retardation of one of the two orthogonal components of the optical field with respect to the other. In general, a beam of light incident on such a material will bifurcate into two beams with two distinct refractive indices and polarization states. The two polarization states corresponding to a specified wave direction in a birefringent material are called the *ordinary* and the *extraordinary states*. *Dichroism* is the material property that results in the difference in attenuation between two orthogonal polarization states. Both of these effects may be present simultaneously. Quartz and calcite are the most commonly used crystals for retarders and prism polarizers. Because of its weak linear birefringence (i.e.,  $\Delta n = 0.0091$  for  $0.5892\text{ }\mu\text{m}$ ), quartz is useful for retardation plates. Prism polarizers are usually made of calcite, because the transmission and birefringence are large, ranging from  $0.200$  to  $1.200\text{ }\mu\text{m}$  (e.g.,  $\Delta n = -0.172$  for  $0.5892\text{ }\mu\text{m}$  [4, 5]).

The characteristics of optical polarizing components have been studied by various authors. Here, we mention in particular the review paper by Bennett

[22]. This chapter is intended only to summarize some of the information that an optical systems designer needs to consider before deciding what sort of optical elements to adopt for performing a desired transformation on light incident on the system. This review is by no means complete but merely identifies the major optical components we have dealt with in the previous chapters.

The order of presentation is as follows. We start in Section 5.2.1 by briefly discussing, various kinds of conventional polarizers. Here we describe reflection, transmission, and selective absorption polarizers. Next, Section 5.1.2 is concerned with compensators.

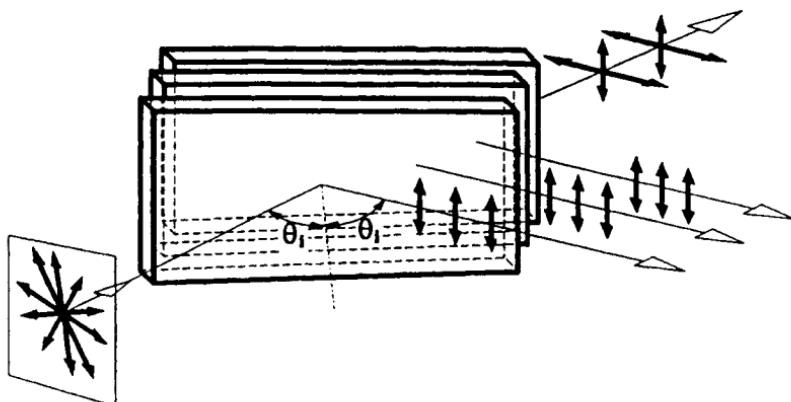
## 5.2.1. POLARIZERS

A *polarizer* is a device that changes the amplitudes of the orthogonal components of the field vector inequally. In some polarizers (e.g., wire-grid polarizer, Polaroid), one of these components is extinguished. In others (e.g. Wollaston prism), both perpendicularly polarized beams at the output of the beam splitter can be measured simultaneously. There are three common types of polarizers: polarizers by reflection, by transmission, and selective absorption. Two qualities of a good polarizer are that gain should very closely equal unity, and the transmitted beam should be uniform in its polarization orientation. Imperfections of the surface or defects within the crystal can scatter the light, allowing multiple optical paths that interfere and consequently can have some depolarizing tendency for incident completely polarized light.

In this section we deal only with linear polarizers. Circular polarizers are also important, for example, in glare-reduction screens for computer monitors. By definition, a circular polarizer extinguishes circularly polarized light of one sense (right- or left-handed), but allows the other circular polarization to pass. As explained in Section 4.1, a series combination of an appropriately oriented linear polarizer and a quarter-waveplate will perform as a circular polarizer. For a detailed discussion of the merits of each of these devices and of their ranges of angle of incidence and wavelength, we refer to Bennett [22] and Shurcliff [23].

### 5.2.1.1. Reflection Polarizers

We have seen in Section 4.2 that whenever unpolarized light is reflected obliquely at the surface of dielectric matter, the reflected beam is partially linearly polarized [i.e., Eq. (4.2.4)]. According to Fresnel's equations, the refracted beam is also partially polarized. It is interesting to observe that the term *reflection polarizer* is used irrespective of whether the reflected beam or the transmitted beam is the one that is retained and used [23]. The design of a reflection polarizer can be done using a single reflecting surface or a pile-of-plates configuration (Fig. 5.2.1). The degree of polarization as a function



**Figure 5.2.1.** The pile-of-plates polarizer.

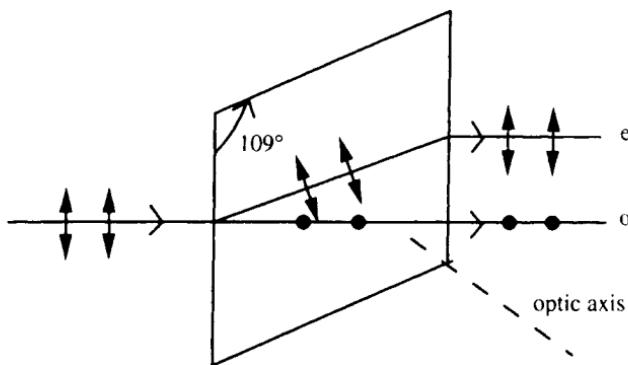
of the incident angle  $\theta_i$  for incident unpolarized light on a pile of  $N$  glass plates has been computed in Section 4.2; see Fig. 4.2.2. When the number of plates within the pile is large, quite high degrees of polarization can be achieved.

Also, on the basis of light reflection to polarize light, intracavity Brewster angle are windows used in laser technology to produce linearly polarized light [24].

### 5.2.1.2. Transmission Polarizers

We now consider linear birefringent beamsplitters. There is a bewildering variety of linear polarizing beamsplitters, including the Glan–Foucault, the Glan–Thompson, the Rochon, the Nicol, the Wollaston, the Senarmont, and the MacNeille prisms. In the visible spectrum, prism polarizers have the highest polarization resolution. An ideal polarizing birefringent beamsplitter is a biprism that separates the ordinary and extraordinary beams by introducing a discontinuity in the refractive index. Figure 5.2.2 shows what happens to the incident linearly polarized beam passing through a birefringent polarizer, specifically, a plane parallel calcite plate. If the crystal is a plane and parallel plate, and the optical axis direction is not parallel with the beam, the light will emerge as two separate, orthogonally polarized beams. Observe that the extraordinary beam emerges laterally displaced. We note also that there are other birefringent prisms, including the Nicol and Glan–Foucault prisms, besides the type depicted in Fig. 5.2.2 for which the ordinary beam undergoes total internal reflection and is absorbed in black paint on the sides of the prism. Thus, only the extraordinary beam is transmitted. Ideally, the transmitted component is completely polarized and suffers no decrease in intensity.

Among the most used polarizing beamsplitters is the Wollaston prism. The incident beam strikes the surface of the prism normally. Figure 5.2.3 shows the ordinary and extraordinary beams for a Wollaston. A Wollaston polarizing

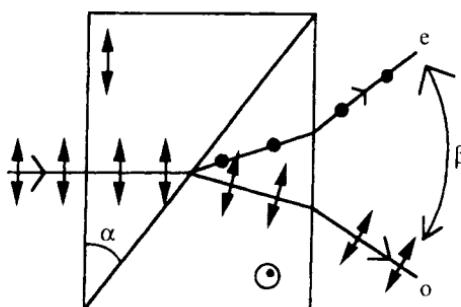


**Figure 5.2.2.** Birefringent polarizer. The light incident on the crystal is split into two linearly polarized beams with perpendicular polarization directions. The symbols *e* and *o* respectively denote the extraordinary and ordinary beams. The extraordinary beam is deviated, while the ordinary beam is not. The direction of the optical axis is indicated by the dashed line.

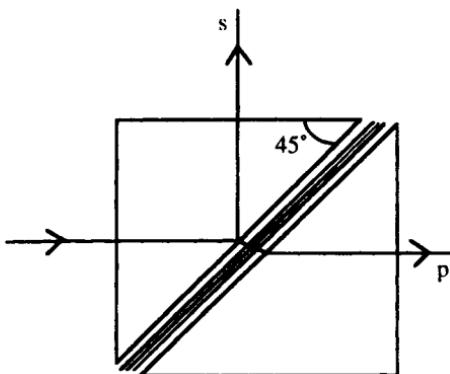
beamsplitter has the optical axis of its first prism parallel to the face to produce ordinary and extraordinary beams that see the second prism's optical axis rotated  $\pi/2$  about the beam axis to produce a maximum birefringence and deviation for both beams.

The Wollaston prism gives both perpendicularly polarized beams with at least 99.9% polarization and large angular separation (i.e., up to  $45^\circ$ ) between the two beams. Denoting by  $\alpha$  the angle of the outer wedges, the angle  $\beta$  between the two emerging orthogonally polarized beam is

$$\beta \cong 2 \tan^{-1}(|\Delta n| \tan(\alpha)) \quad (5.2.1)$$



**Figure 5.2.3.** Wollaston prism made of two wedges of a uniaxial crystal with optic axes perpendicular to the direction of light. The optical axis direction is oriented parallel ( $\uparrow$ ) or perpendicular to ( $\odot$ ) to the page. The symbol  $\beta$  denotes the angular separation;  $\alpha$  is the construction angle of the outer wedges.



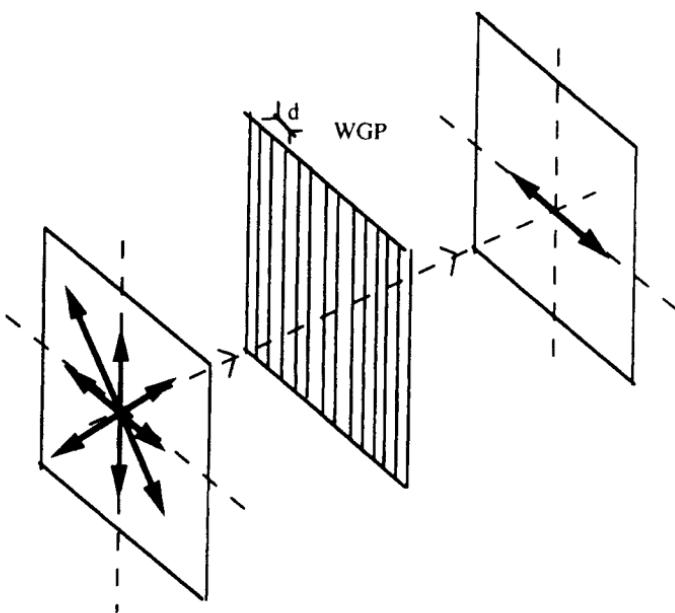
**Figure 5.2.4.** The MacNeille polarizing beamsplitter cube.

where  $\Delta n$  is the prism birefringence [26]. This device has good performance in applications where two separate beams of orthogonal polarizations are required.

A thin-film polarizing beamsplitter cube divides the incident beam into s (perpendicular) and p (parallel) polarization states, transmitting the p-polarized component (horizontal linearly polarized Stokes vector  $([1 \ 1 \ 0 \ 0]^T)$ ) and reflecting the s-polarized component (vertical linearly polarized Stokes vector  $([1 \ -1 \ 0 \ 0]^T)$ ). For example, the MacNeille polarizing beamsplitter is constructed by coating the hypotenuse face of an isosceles right-angled prism with multilayered thin films (alternating layers of materials of high and low refraction indices) and then cementing it to an identical prism (see Fig. 5.2.4). The thickness of each layer and the refractive indices of the thin-film materials are such that the cube acts as a polarizer for the s component in reflection and the p component in transmission [25–27].

### 5.2.1.3. Polarization by Selective Absorption

An efficient transmission polarizer in the infrared and submillimeter spectral ranges is a grid of parallel wires. Historically it is worth noting that Heinrich Hertz used such grids to test the properties of radiowaves. A *wire-grid polarizer* (WGP) is composed of an infrared (IR)-transparent substrate on which parallel lines of conducting material have been deposited. Suppose that an unpolarized wave is incident on the grid, as illustrated in Figure 5.2.5. As discussed in Part 3, unpolarized light can be decomposed in a sum of two orthogonal pure states, Eq. (3.1.163). The response of the WGP varies according to each elementary pure component. The polarization parallel to the lines is absorbed [the electric field parallel component of the lightwave drives electrons in the wires, thus generating a current within the lines that causes resistive (Joule) heating and thereby loses energy] while the polarization perpendicular to the lines is



**Figure 5.2.5.** Schematic of the wire-grid polarizer (WGP). Polarization by selective absorption. The grid spacing is  $d$ . The gain for unpolarized light passing through an ideal polarizer is constantly equal to one-half, regardless of the orientation of the wire-grid polarizer. The gain for a partially polarized wave in general is sinusoidal as a function of polarizer orientation oscillating between a minimum and a maximum every  $\pi/2$  [see Eq. (4.1.36)].

transmitted (it has no electrons to drive and hence passes through without much loss). Thus the transmission axis of the grid is perpendicular to the wires. If the spacing  $d$  between the wires is about five times larger than the diameter of the wires, the polarizer is efficient for wavelengths longer than  $5d$ . There are efficient WGP's in the  $3\text{--}30\ \mu\text{m}$  spectral region, and their gain for unpolarized light  $g_u$  is  $\cong 0.2$  at  $3\ \mu\text{m}$  and  $\cong 0.4$  at wavelengths longer than  $8\ \mu\text{m}$  [29].

The dichroic polarizing sheets, commonly called *Polaroids*, are molecular analogs of the wire grid. The development of these optical systems is due largely to the pioneering researches of Land [30]. Their applications in optics have established their status as the most widely used linear polarizers. These efficient polarizing materials are made of a flexible, thin polymer film, polyvinyl alcohol, which has been stretched in a given direction so as to align their long-chain hydrocarbon molecules and then impregnated with a solution containing dichroic species such as iodine. The iodine attaches to the polymer chains, forming a chain of its own. The component of the electric field of an incident light that is parallel to the molecules drives the electrons and, similar to the wire grid, is strongly absorbed. Thus the transmission axis of the

Polaroid is perpendicular to the direction in which the film was stretched. A wide variety of sheet-type polarizers are available. The early history of these devices has been reviewed by Shurcliff [23].

### 5.2.2. COMPENSATORS

Ideal compensators are optical devices that introduce a differential phase shift between the linear orthogonal components of an incident light, with no accompanying intensity attenuation. The usual compensator used in polarimetry instruments is a plane parallel plate of uniaxial or biaxial material, such as quartz and mica, with its optical axis oriented parallel to the entrance surface. Other conventional compensators are based on the difference between the p- and s-reflection phase shifts in total internal reflection. For a review of the methods for phase retardation measurements in birefringent anisotropic crystal plates with a discussion of their sensitivity, the reader may wish to consult Wood and Glazer [31] and Title [32].

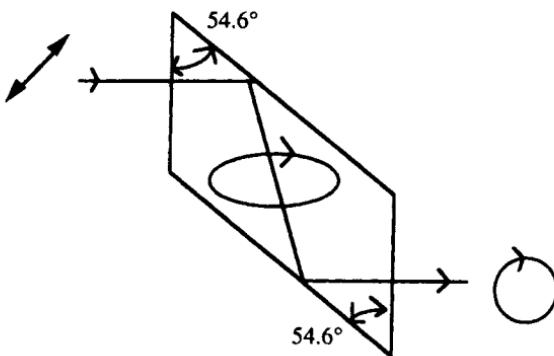
If the thickness  $L$  of the plate is such that the differential phase shift between the two orthogonal components 1 and 2 (see Section 4.1.2.3) is  $\lambda/4$ , where  $\lambda$  is the wavelength in vacuum, the plate is called a *quarter-waveplate*; it changes a circularly polarized wave into a linearly polarized wave. If the thickness of the plate is such that the differential phase shift is  $\lambda/2$ , the plate is called a *half-waveplate*; it changes right-handed circularly polarized waves into left-handed.

Note that as the refractive indices of the crystal for the ordinary,  $n_o$ , and extraordinary  $n_e$ , beams, are wavelength-dependent, the phase shift, for normal incident radiation, between linear orthogonal polarizations  $\delta = 2\pi(L/\lambda)|\Delta n|$  is also wavelength-dependent. The quantity  $\delta\lambda/2\pi$  is called the *pathlength difference*. A fast and slow coordinate system oriented with respect to the optical axis characterizes the compensator's orientation. For  $\Delta n > 0$ , linearly polarized light travels faster parallel to the optical axis and for  $\Delta n < 0$ , the fast direction is perpendicular to the optical axis.

Compensators are usually made of quartz, mica, or plastic. Of the many different kinds of compensators, we shall consider only two of those that are commonly used: the Fresnel rhomb and the Babinet compensator. If a linear polarizer is combined with an adjustable compensator, any type of polarization can be made (i.e., circular and elliptical states). Variable-retardation plates are used also to modulate the different Mueller matrix elements onto intensity variations at separate frequencies. Finally we note that magnetically controllable waveplates using ferrofluids have been designed for fast polarimetric modulation.

#### 5.2.2.1. Fresnel Rhomb

Let us consider the geometry shown in Fig. 5.2.6. A linearly polarized wave whose direction of polarization is at an angle of  $45^\circ$  with respect to the face edge of the rhomb of glass is normally incident on one face. The light beam in



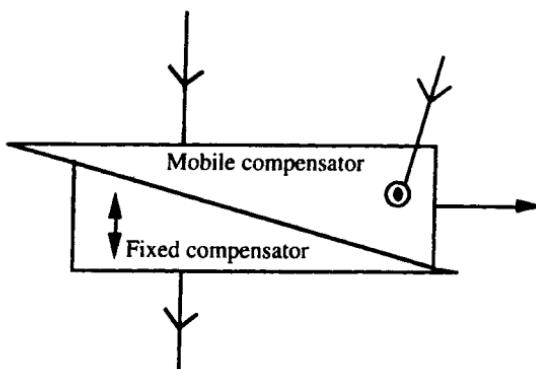
**Figure 5.2.6.** The Fresnel rhomb is a glass prism ( $n = 1.51$ ) made in the form of a rhomb having apex angles of  $54.6^\circ$ .

the Fresnel rhomb undergoes two total internal reflections before emerging from the exit face, resulting in a quarter-wavepath wave difference. Thus the emerging wave is circularly polarized. If the index is 1.51, the angle of each apex of the rhomb should be  $54.6'$ .

On one hand, Fresnel rhombs operate properly over a wide range of wavelengths, but on the other hand they have the disadvantages of being large in construction and introducing a displacement of the light beam. In many instances, a combination of laminate compensators may be more convenient.

### 5.2.2.2. Adjustable Compensators

The Babinet compensator is a continuously adjustable compensator. This device changes the thickness, experiencing a birefringence, which in turn,



**Figure 5.2.7.** Principle of the Babinet compensator. It consists of a fixed wedge and a movable wedge with the optical axis parallel to the entrance and exit surfaces, but crossed to each other.

changes the phase difference between the two linear orthogonal polarizations. The principle of this optical device is as follows. Let us consider a pair of crystalline wedges whose optical axes are perpendicular to each other, as shown in Fig. 5.2.7. The roles of the ordinary and extraordinary beams are interchanged as light passes from the upper to the lower wedges. Because of this interchange, any phase shift that may have accumulated in one wedge can be canceled by the other. If the two wedges can move relative to one another, we can get any desired phase shift  $\delta$ . Observe that  $\delta$  will vary from point to point over the surface.

Variable compensators can also be achieved using elastooptic, electrooptic and magnetooptic effects [2, 5]. Electrooptic modulators are based on either the linear Pockels effect or the quadratic Kerr effect. Magnetooptic effects are the Faraday, Kerr, and Cotton–Mouton effects. Discussion of these effects is beyond the scope of this book; interested readers are referred to Refs. 4 and 5.

With this material on optical polarizing components, we can now turn to the study of simple approaches for measuring the Stokes parameters and the polarization matrices.

## SECTION 5.3

# Measurement of Stokes Parameters

The polarization of light is fully described by the four elements of the Stokes vector, the measurement of which has become possible because of advances in experimental technology. The statistics of the Stokes parameters is a useful tool for studying scattering media that modify the polarization state of light. Much of the current work on polarization optics utilizes optical techniques that have been developed in the past few decades. Areas of application include characterization of light sources, measurement of the polarization-altering properties of optical systems, and astronomical measurements, to name only a few. A growing body of methodologies has been developed to completely characterize the polarization of any light beam. Here we present a sampling of some of the techniques that are used to analyze polarization in transmitted light by an optical component, or in scattered light.

The primary purpose of Section 5.3.1 is to describe several procedures for measuring the Stokes parameters. Next, in Section 5.3.2, we shall briefly discuss measurement methods of the probability density function of the Stokes parameters.

### 5.3.1. METHODS FOR MEASURING STOKES PARAMETERS

In this section, we will consider quasimonochromatic light that is incident on a detector whose response time,  $\tau_3$ , is much larger than the time of coherence,  $\tau_2$ . From the discussion in Part 3, it follows that the light is, in general, partially polarized. Because they are the observables of light at optical frequencies, the time-averaged Stokes parameters are directly accessible to measurement. Several schemes have evolved over the years to evaluate the four elements of the Stokes vector. These schemes exploit one or more of the changes that transmission through a train of compensators, polarizers, and rotators induces on the intensity of an incident light. Some methods measure the four components of the Stokes vector separately, while others measure them simultaneously. It should be noted that appraisals of the merits of a measurement technique are always relative to its alternatives. For details regarding accuracy, comparisons, and amenability to computer control, we refer the reader to Refs. 33–40.

### 5.3.1.1. Stokes' Procedure

In the Stokes scheme, the measurement of the state of polarization of light is performed by rotating a retarding waveplate and a polarizer to four pairs of angular orientations and recording the four transmittances through the polarizer. Figure 5.3.1 shows a schematic of the multiple-component serial device that points out the important polarization elements. A beam of light from the quasimonochromatic source  $S$  is incident on a rotating compensator of Mueller matrix  $C(\delta)$  followed by a linear polarizer of Mueller matrix  $P(\alpha)$ . The emerging beam of light finally reaches an optical detector  $D$ , such as a photodiode or a photomultiplier tube, which is assumed to be polarization insensitive.

We denote by  $\langle S_j(\alpha, \delta) \rangle$  the averaged intensity of light, at the output of the linear analyzer  $P(\alpha)$  with its electric field vector in a direction at an angle  $\alpha$  to the 1-axis when the 2-axis component of the electric field vector is subjected to a retardation  $\delta$  ( $\delta \ll (\Delta\nu/v_0)$ ) with respect to the 1-axis component. From Eq. (4.1.87) and making use of the Mueller matrices of these optical components, we may readily deduce that this intensity is given by

$$\langle S_0(\alpha, \delta) \rangle = \frac{1}{2}(\langle S_0 \rangle + \cos(2\alpha)\langle S_1 \rangle + \sin(2\alpha)(\langle S_2 \rangle \cos(\delta) + \langle S_3 \rangle \sin(\delta))) \quad (5.3.1)$$

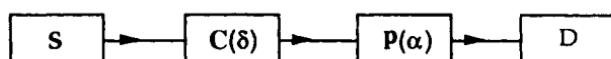
Equation (5.3.1) is saying that the Stokes parameters may be determined by the measurements for a number of selected values of  $\delta$  and  $\alpha$ . These are found from Eq. (5.3.1) to be

$$\langle S_0 \rangle = \langle S_0(0, 0) \rangle + \left\langle S_0\left(\frac{\pi}{2}, 0\right) \right\rangle \quad (5.3.2a)$$

$$\langle S_1 \rangle = \langle S_0(0, 0) \rangle - \left\langle S_0\left(\frac{\pi}{2}, 0\right) \right\rangle \quad (5.3.2b)$$

$$\langle S_2 \rangle = \left\langle S_0\left(\frac{\pi}{4}, 0\right) \right\rangle - \left\langle S_0\left(\frac{3\pi}{4}, 0\right) \right\rangle = 2\left\langle S_0\left(\frac{\pi}{4}, 0\right) \right\rangle - \langle S_0 \rangle \quad (5.3.2c)$$

$$\langle S_3 \rangle = \left\langle S_0\left(\frac{\pi}{4}, \frac{\pi}{2}\right) \right\rangle - \left\langle S_0\left(\frac{3\pi}{4}, \frac{\pi}{2}\right) \right\rangle = \langle S_0 \rangle - 2\left\langle S_0\left(\frac{\pi}{4}, \frac{\pi}{2}\right) \right\rangle \quad (5.3.2d)$$



**Figure 5.3.1.** The different components of the experimental setup: the input light source  $S$  (e.g., HeNe laser), the compensator  $C(\delta)$ , the linear polarizer  $P(\alpha)$ , and the radiation detector  $D$ .

The crucial steps in the determination of the four Stokes parameters may be summarized as follows:

1. Set up the configuration displayed in Fig. 5.3.1.
2. Measure the intensity of the light transmitted through the device for selected values of  $\delta$  and  $\alpha$ . The first three parameters are measured by removing the compensator and setting the angle of transmission axis of linear polarizer to 0,  $\pi/4$ , and  $\pi/2$ . The last Stokes parameter is obtained by reinserting the compensator and rotating the angle of the linear polarizer to  $\pi/4$ .
3. Calculate the Stokes parameters given by Eqs. (5.3.2a–d).

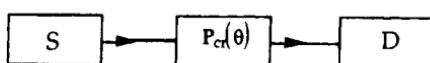
According to Eq. (5.3.1), it follows simply that the degree of polarization of the incident wave obeys the equation

$$\frac{g_{\max} - g_{\min}}{g_{\max} + g_{\min}} = P_i \quad (5.3.3)$$

The errors associated with this method of measurement may be due to the misalignment of the fast axis of the compensator, the internal properties (absorption) of the birefringent material, and the partially polarizing detector. To eliminate the errors due to the compensator in the measurement of  $\langle S_3 \rangle$ , an alternative approach was proposed by Collett [40].

### 5.3.1.2. Collett's Procedure

This method simplifies the Stokes procedure by eliminating the use of a compensator in the procedure. It requires only a single circular polarizer, constructed from a quarter-wave retardation plate and a linear polarizer whose transmission axis is at  $45^\circ$  with respect to the horizontal axis, placed in a rotatable mechanical mount. The Stokes parameters are measured at selected values of the rotation angle. The circular polarizer is then flipped  $180^\circ$ , and the Stokes parameters are again measured at selected values of the rotation angle. Figure 5.3.2 shows the optical components that are used in the measurement apparatus. The Mueller matrix for the rotating right circular polarizer is



**Figure 5.3.2.** Schematic diagram of the apparatus used in measuring the Stokes parameters according to Collett's procedure:  $P_{cr}$  is a rotatable right circular polarizer;  $S$  is the input light source.

expressed as

$$\mathbf{P}_{cr}(\theta) = \mathbf{R}(-\theta)\mathbf{C}(0)\mathbf{P}\left(\frac{\pi}{4}\right)\mathbf{R}(\theta) \quad (5.3.4)$$

where  $\mathbf{R}(\theta)$  is the Mueller matrix of a rotator corresponding to a rotation of an angle  $\theta$ , Eq. (4.1.116). Consequently the Mueller matrix for the rotating circular polarizer reads as

$$\mathbf{P}_{cr}(\theta) = \frac{1}{2} \begin{bmatrix} 1 & -\sin(2\theta) & \cos(2\theta) & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -\sin(2\theta) & \cos(2\theta) & 0 \end{bmatrix} \quad (5.3.5)$$

The output-averaged intensity emerging from the circular polarizer is

$$\langle S_0(\theta) \rangle_c = \frac{1}{2} (\langle S_0 \rangle - \sin(2\theta)\langle S_1 \rangle + \langle S_2 \rangle \cos(2\theta)) \quad (5.3.6)$$

where the subscript “c” refers to the fact that describes the circular side of the polarizer. Note that Eq. (5.3.7) depends only on the Stokes parameters for linear states. The Mueller matrix when the circular polarizer is flipped to its linear side is

$$\mathbf{P}_l(\theta) = \mathbf{R}(-\theta)\mathbf{P}(0)\mathbf{R}(\theta) = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\ \sin(2\theta) & 0 & 0 & \sin(2\theta) \\ -\cos(2\theta) & 0 & 0 & -\cos(2\theta) \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (5.3.7)$$

In like fashion, the output averaged intensity emerging from the linear polarizer is

$$\langle S_0(\theta) \rangle_l = \frac{1}{2} (\langle S_0 \rangle + \langle S_3 \rangle) \quad (5.3.8)$$

where the subscript l refers to the fact that Eq. (5.3.7) describes the linear side of the polarizer combination. Now the intensity depends only of the circulary polarized component of the input light beam, independent of the rotation angle. Thus the formulas for the Stokes parameters are

$$\langle S_0 \rangle = \langle S_0(0) \rangle_c + \left\langle S_0\left(\frac{\pi}{2}\right) \right\rangle_c \quad (5.3.9a)$$

$$\langle S_1 \rangle = \langle S_0 \rangle - 2 \left\langle S_0\left(\frac{\pi}{4}\right) \right\rangle_c \quad (5.3.9b)$$

$$\langle S_2 \rangle = \langle S_0(0) \rangle - \left\langle S_0 \left( \frac{\pi}{2} \right) \right\rangle_c \quad (5.3.9c)$$

$$\langle S_3 \rangle = 2\langle S_0(0) \rangle_1 - \langle S_0 \rangle \quad (5.3.9d)$$

Recapitulation of the algorithm: the various steps can be summarized as follows:

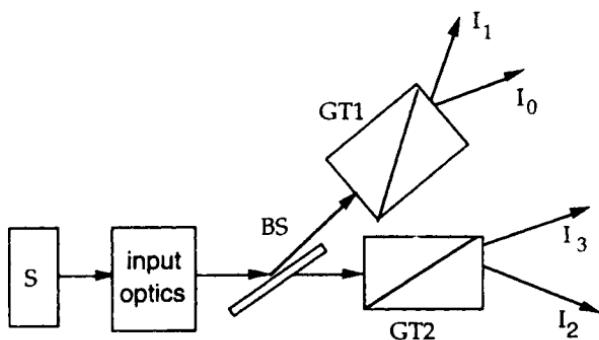
1. Set up the configuration displayed in Fig. 5.3.2.
2. Measure the intensity of the light transmitted through the device for selected values of the rotation angle (i.e.,  $\theta = 0, \pi/2, \pi/4$ ), then invert the right-circular polarizer and measure the intensity of the light transmitted.
3. Calculate the Stokes parameters by using Eqs. (5.3.9a–d).

The most significant error associated with this procedure comes from the polarizer. For a further detailed discussion of the calibration scheme, I refer the interested reader to Ref. [40]. This arrangement makes it impossible to measure all four Stokes parameters simultaneously. Our attention is now turned to another method of measurement, division-of-amplitude photopolarimetry (DOAP), which is capable of time-resolved simultaneous measurements of the four Stokes parameters. The operating principle for this polarimeter and its use for measuring the Stokes parameters is reviewed below.

### 5.3.1.3. Division-of-Amplitude Photopolarimetry

We now consider the technique based on DOAP. The development of this method lean heavily on the research efforts of Azzam [42]. This measurement procedure is rapid and accurate; the mean deviations of the Stokes parameters are on the order of percent. For details concerning the construction and the calibration of the DOAP apparatus, we refer to Refs. [42 and 43]. Figure 5.3.3 shows a block diagram of the apparatus configured to measure the incoming Stokes vector. The principle of DOAP method is as follows. The light beam is impinging on a special coated beamsplitter BS (it requires an eight-wave optical thickness of a low-index material on a high-refractive-index substrate), which transmits and reflects light into a pair of identical beamsplitting Glan–Thompson (or, equivalently, Wollaston prisms) polarizers oriented at  $45^\circ$  with respect to the reference plane. No other optical elements are needed. The four beams from the two Glan–Thompson polarizers are detected by four photodetectors that produce intensities  $I_0, I_1, I_2$ , and  $I_3$ . In this stationary arrangement, the relative placement of the four detectors is crucial.

Denoting the incoming Stokes vector by  $S_i$ , the output Stokes vector of the apparatus is  $S_o = MS_i$ , where  $M$  is the Mueller matrix of the DOAP at a given frequency. To evaluate the elements of the Mueller matrix, the algorithm



**Figure 5.3.3.** Schematic representation of the DOAP apparatus under consideration.

proceeds as follows:

1. Calibrate the instrument with four linearly independent incident Stokes vectors to measure the instrument matrix  $\mathbf{M}$ .
2. Measure the output states vector for selected states of polarizations.
3. Invert  $\mathbf{S}_i = \mathbf{M}^{-1}\mathbf{S}_o$ . Azzam and coworkers have shown that the essential condition that the instrument matrix  $\mathbf{M}$  be nonsingular (i.e., hence  $\mathbf{M}^{-1}$  exists) requires that the plane of incidence be rotated between successive reflections by other than  $\pi/2$  where  $\pm(\pi/4)$  and  $\pm(3\pi/4)$  are the optimum rotations [42].

This technique is particularly applicable to polarization metrology applications. However, the principle of the measurement is a good deal simpler than in practical application. Experiments using this technique have been performed in investigations of rough heterogeneous and anisotropic surfaces [42, 43].

Our discussion of the methods for measuring the Stokes parameters would be incomplete if we did not refer the reader to additional related developments [33–35].

### 5.3.2. PROBABILITY DENSITY FUNCTIONS OF STOKES PARAMETERS

In the preceding subsection we have described essentially a set of methods commonly employed in polarization optics experiments for measuring the Stokes parameters. In this section, we shall discuss briefly another important subject: the determination of the probability density functions of the Stokes parameters.

Characterization of the temporal fluctuations in the optical field is a nontrivial problem. Nature rarely provides us with an explicit knowledge of

the multivariate joint probability density function (PDF) of the electric field. However, we encountered this problem in Section 3.3, and solutions have already been found. Exact analytic results for the polarization aspect of this problem have been obtained only for the simplest case of a Gaussian optical field. In practice the PDF of the Stokes parameters are of fundamental interest, especially when a second-order theory is not entirely adequate to describe the statistics of the radiation field. In only a few instances have authors characterized the PDF of the Stokes parameters. These issues have become more important with the present trend toward the exploitation of polarization information in speckle patterns.

It is very appropriate to end this section by describing the experimental work in PDF. Some of the research of Fercher and Steeger has contributed to the understanding of the polarization properties of speckle pattern [37]. Another case of practical interest was considered by Freund and concerned the measurement of the PDF of  $S_0$  by videopolarimetry [38]. The experiments were carried out in reflection of a HeNe (helium-neon) laser beam by different scattering media, and involved videographic recording of the appropriate speckle patterns. As far as I know, the corresponding PDFs of the other temporal Stokes parameters have not been fully investigated.

## **SECTION 5.4**

---

# **Measurement of the Jones and Mueller Polarization Matrices**

This chapter is devoted to presenting and discussing a number of available techniques that have been developed for measuring the Jones and Mueller polarization matrices that characterize optical elements. These matrices, which show mathematically how these optical elements affect the polarization of a light beam, have achieved a currency among not only optical physicists but also physical chemists studying the interaction of partially polarized light with complex fluids [42, 43]. The common element of these problems is the linearity of the relations involved. We consider a few of them in the following sections, with a view toward illustrating some of their measurement methods as well as some of the analytic tools used in the development of these methods. A key feature of all these developments is the desire to determine the polarization matrix elements with the best available accuracy. The subject is very extensive and has experienced a wide variety of applications. These include fields as diverse as oceanography, biophysics, and atmospheric physics, to mention only a few.

We first concentrate on the measurement procedures of the Jones matrix elements, in Section 5.4.1. We review a set of methods that have been developed over the years by a number of investigators. Unnecessary details are avoided and the material is presented from an applications-oriented viewpoint so that the reader will hopefully assimilate a practical working knowledge of the subject. More specifically, we try to highlight details of their algorithmic implementation. The next section extends this presentation of several measurement methods of all 16 elements of the Mueller polarization matrices, including the polarization-modulation technique. The thrust of this section is on the underlying physical ideas and not on applications of these concepts to technologically useful applications. It is important to emphasize at the outset that one must resort to a suitable coordinate system to ensure that each optical element can be mathematically represented by its appropriate polarization matrix.

### 5.4.1. METHODS OF DETERMINATION OF JONES MATRIX

There are at least two classic routes for determining a Jones matrix. The first has been developed by Clark Jones [44]. The second has found widespread application in polarization spectroscopy [45–47].

#### 5.4.1.1. Jones Procedure

In his pioneering researches, Clark Jones addressed the important problem of measuring the four matrix elements of  $\mathbf{J}$  [44]. More specifically, in the seventh paper of his series, he introduced a measurement procedure to determine the four complex elements of the Jones matrix: to extract physical information from the transmission characteristics. The main purpose of this section is to discuss in detail the basic theory and implementation of the Jones derivation. The coordinate is chosen so that the light is propagating in the  $x_3$  direction. The various steps of the Jones algorithm can be formally stated as follows:

1. Use incident horizontally polarized light and detect the emerging elliptically polarized light. The emerging Jones vector is given by

$$\begin{bmatrix} J_{11} \\ J_{21} \end{bmatrix}$$

For later purposes, we set  $a_1 = J_{21}/J_{11}$ .

2. Repeat step 1, but with an incident vertically polarized light, and detect the emerging elliptically polarized light. The output Jones vector now reads

$$\begin{bmatrix} J_{12} \\ J_{22} \end{bmatrix}$$

and we define  $a_2 = J_{12}/J_{22}$ .

3. Repeat step 1, but with the reversed direction of propagation of light. The emerging Jones vector is

$$\begin{bmatrix} J_{11} \\ J_{22} \end{bmatrix}$$

and we define  $a_3 = J_{22}/J_{11}$ .

4. Similarly, repeat step 2, but with the reversed direction of propagation of light. The output Jones vector writes

$$\begin{bmatrix} J_{12} \\ J_{21} \end{bmatrix}$$

and we define  $a_4 = J_{12}/J_{21}$ . It follows that

$$a_1 a_4 = a_2 a_3 \quad (5.4.1)$$

In other words, this is not an independent determination. Thus the Jones matrix has the form

$$\mathbf{J} = a \begin{bmatrix} 1 & a_3 \\ a_1 & a_2 a_3 \end{bmatrix} \quad (5.4.2)$$

where  $a \equiv |a| \exp(i\phi)$ .

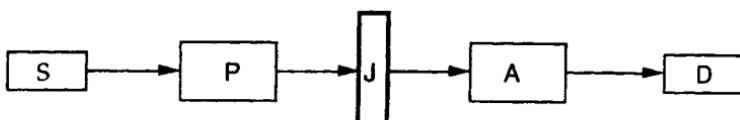
5. Finally we see from Eq. (4.1.24) that  $|a| = [2g_u]^{1/2}/(1 + |a_1|^2 + |a_3|^2 + |a_2 a_3|^2)^{1/2}$ , where  $g_u$  is the gain for an incident unpolarized light. It follows that the Jones matrix  $\mathbf{J}$  is determined up to overall (non-measurable) phase  $\phi$ .

#### 5.4.1.2. Differential Polarized Spectroscopy

The formulas of the gain that we derived in Section 5.4 may be applied in a straightforward manner to determine the Jones matrix for an unknown device. The method is based on a set of experiments in which known polarization states are sent into the optical system and the corresponding output states are measured. With this remark, we now outline a derivation of the differential polarized spectroscopy (DPS) originally due to Raab [45]. Reference is made to Brosseau for an application of this method [47].

The concept of DPS is shown in the block diagram of Fig. 5.4.1. The arrangement can be divided into three sections: the state of polarization generator, the sample, and the state of polarization analyzer. In such multiple-component serial device, a well-collimated quasimonochromatic beam from a light source  $S$  passes through polarizing optics  $P$  to produce light of known controlled polarization. These polarization states interact with the optical medium  $\mathbf{J}$  under study and are modified. The output state of polarization is further characterized by the analyzing optics  $A$ . The photodetector  $D$  is assumed to be isotropic and recording of intensity is done regardless of polarization. We assume that all optical components are perfect, with no reflection losses, for example.

The general procedure consists in launching a wave of unit intensity in the form of a pure state in one of the six basic polarization states (linearly polarized horizontally, linearly polarized vertically, linearly polarized at  $+45^\circ$ , linearly polarized at  $-45^\circ$ ; circularly polarized right, and circularly polarized left) and to analyze each transmitted beam with respect to these six reference states. This gives a total of 36 analyzed intensities, which are expressible in terms of the elements of the Jones matrix  $\mathbf{J}$ . Because of the linearity,



**Figure 5.4.1.** The different components of the typical experimental setup to measure the elements of the Jones matrix: the light source  $S$ ; a polarizing unit  $P$  to prepare the incident state in one of the six basic polarization forms—linearly polarized horizontally, linearly polarized vertically, linearly polarized at  $+45^\circ$ , linearly polarized at  $-45^\circ$ , and circularly polarized right and circularly polarized left; the optical medium characterized by a Jones matrix  $J$ ; the analyzing unit  $A$  to characterize the transmitted beam with respect to the six polarization forms; and the radiation detector  $D$ .

characterization of the input and output states of polarization repeated for a limited number of different input and analyzed states leads to the determination of the Jones matrix components, namely, the law of transformation of all input polarization states by the optical medium. The basis of DPS is to measure the difference between two transmitted intensities analyzed for particular polarization states, when two input beams of same intensity, each of a given polarization state, are alternatively incident on the optical medium.

In general, a Jones matrix has eight real matrix elements corresponding to the eight elementary dielectric properties. DPS permits evaluation of the contributions of each type of birefringence and dichroism, even if many of these effects are exhibited simultaneously. We find it convenient to denote by  $(i, a)$  the analyzed intensity that is measured when a unit intensity of pure state  $i$  is incident and the transmitted beam is analyzed for particular state of polarization  $a$ . The difference between two such intensities will be denoted by  $[(i, a) - (i', a')]$ . It is important to appreciate that the measurements of intensities and difference of intensities permit evaluation of the values of  $|J_{11}|^2$ ,  $|J_{12}|^2$ ,  $|J_{21}|^2$ ,  $|J_{22}|^2$ ,  $\text{Re}(J_{11}J_{22}^*)$ ,  $\text{Re}(J_{11}J_{21}^*)$ ,  $\text{Re}(J_{12}J_{21}^*)$ ,  $\text{Re}(J_{12}J_{22}^*)$ ,  $\text{Im}(J_{11}J_{22}^*)$ ,  $\text{Im}(J_{11}J_{21}^*)$ ,  $\text{Im}(J_{12}J_{21}^*)$ ,  $\text{Im}(J_{11}J_{12}^*)$ ,  $\text{Im}(J_{12}J_{21}^*)$ , and  $\text{Im}(J_{21}J_{22}^*)$ . We incorporated the results of such calculations for all 36 combinations into the matrix array shown in Table 5.4.3. We adopt a notation in which both the incident and analyzed states of polarization are displayed. For example, we see that  $\text{Re}(J_{11}J_{12}^*) = \frac{1}{2}[(\nearrow, \leftrightarrow) - (\nwarrow, \leftrightarrow)]$ . Thus the element  $\text{Re}(J_{11}J_{12}^*)$  is measured by illuminating the Jones medium with  $\pm(\pi/4)$  linear polarized light while analyzing the output light with a linear horizontal polarizer.

This matrix array contains several things worth mentioning. In general, all methods for birefringence measurements are also sensitive to linear dichroism. We illustrate the potential of this method by considering a simple example. Our illustrating example is confined to the case of a monoclinic crystals oriented in a plane (010) that exhibits linear birefringence and linear dichroism

TABLE 5.4.1. A Matrix Array Showing the Jones Matrix Element Combination Measured for Various Arrangements of Input and Analyzed States of Polarization<sup>a</sup>

Incident Analyzed	$\leftrightarrow$	$\uparrow$	$\swarrow$	$\nearrow$	$\odot$	$\circlearrowleft$
$\leftrightarrow$	$ J_{11} ^2$	$ J_{12} ^2$	$A + \text{Re}(J_{11}J_{12}^*)$	$A - \text{Re}(J_{11}J_{12}^*)$	$A - \text{Re}(J_{11}J_{12}^*)$	$A + \text{Re}(J_{11}J_{12}^*)$
$\uparrow$	$ J_{21} ^2$	$ J_{22} ^2$	$B + \text{Re}(J_{21}J_{22}^*)$	$B - \text{Re}(J_{21}J_{22}^*)$	$B - \text{Im}(J_{21}J_{22}^*)$	$B + \text{Im}(J_{21}J_{22}^*)$
$\swarrow$	$C + \text{Re}(J_{11}J_{21}^*)$	$D + \text{Re}(J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{11}J_{21}^* + J_{11}J_{22}^*$ $+ J_{12}J_{21}^* + J_{12}J_{22}^*$ $+ J_{21}J_{22}^*$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{11}J_{21}^* + J_{11}J_{22}^*$ $+ J_{12}J_{21}^* - J_{12}J_{22}^*$ $+ J_{21}J_{22}^*$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{21}^*)$ $+ J_{12}J_{22}^* - \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{21}J_{22}^*$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{21}^*)$ $- J_{12}J_{22}^* + \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $+ J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{21}J_{22}^*$
$\nearrow$	$C - \text{Re}(J_{11}J_{21}^*)$	$D - \text{Re}(J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{11}J_{21}^* - J_{11}J_{22}^*$ $- J_{12}J_{21}^* - J_{12}J_{22}^*$ $+ J_{21}J_{22}^*$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{11}J_{21}^* - J_{11}J_{22}^*$ $- J_{12}J_{21}^* + J_{12}J_{22}^*$ $+ J_{21}J_{22}^*$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{21}^*)$ $+ J_{12}J_{22}^* - \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $- J_{11}J_{22}^* + J_{12}J_{21}^*$ $+ J_{21}J_{22}^*$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{21}^*)$ $- J_{12}J_{22}^* + \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $- J_{11}J_{22}^* + J_{12}J_{21}^*$ $+ J_{21}J_{22}^*$
$\odot$	$C + \text{Im}(J_{11}J_{21}^*)$	$D + \text{Im}(J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{21}J_{22}^* + \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $+ J_{11}J_{22}^* + J_{12}J_{21}^*$ $+ J_{12}J_{22}^*$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{21}J_{22}^* + \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{12}J_{22}^*$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{22}^*)$ $- J_{12}J_{21}^* - \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{21}J_{22}^*$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{22}^*)$ $- J_{12}J_{21}^* + \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{21}J_{22}^*$
$\circlearrowleft$	$C - \text{Im}(J_{11}J_{21}^*)$	$D - \text{Im}(J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{21}J_{22}^* - \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $+ J_{11}J_{22}^* + J_{12}J_{21}^*$ $+ J_{12}J_{22}^*$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{21}J_{22}^* + \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{12}J_{22}^*$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{22}^*)$ $- J_{12}J_{21}^* - \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $- J_{11}J_{22}^* + J_{12}J_{21}^*$ $+ J_{21}J_{22}^*$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{22}^*)$ $- J_{12}J_{21}^* + \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{21}J_{22}^*$

<sup>a</sup>Here we have set  $A = \frac{1}{2}(|J_{11}|^2 + |J_{12}|^2)$ ,  $B = \frac{1}{2}(|J_{21}|^2 + |J_{22}|^2)$ ,  $C = \frac{1}{2}(|J_{11}|^2 + |J_{21}|^2)$ , and  $D = \frac{1}{2}(|J_{12}|^2 + |J_{22}|^2)$ .

TABLE 5.4.1. A Matrix Array Showing the Jones Matrix Element Combination Measured for Various Arrangements of Input and Analyzed States of Polarization<sup>a</sup>

Incident Analyzed	$\leftrightarrow$	$\uparrow$	$\swarrow$	$\nearrow$	$\odot$	$\odot$
$\leftrightarrow$	$ J_{11} ^2$	$ J_{12} ^2$	$A + \text{Re}(J_{11}J_{12}^*)$	$A - \text{Re}(J_{11}J_{12}^*)$	$A - \text{Re}(J_{11}J_{12}^*)$	$A + \text{Re}(J_{11}J_{12}^*)$
$\uparrow$	$ J_{21} ^2$	$ J_{22} ^2$	$B + \text{Re}(J_{21}J_{22}^*)$	$B - \text{Re}(J_{21}J_{22}^*)$	$B - \text{Im}(J_{21}J_{22}^*)$	$B + \text{Im}(J_{21}J_{22}^*)$
$\swarrow$	$C + \text{Re}(J_{11}J_{21}^*)$	$D + \text{Re}(J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{11}J_{21}^* + J_{11}J_{22}^*$ $+ J_{12}J_{21}^* + J_{12}J_{22}^*$ $+ J_{21}J_{22}^*)$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{11}J_{21}^* + J_{11}J_{22}^*$ $+ J_{12}J_{21}^* - J_{12}J_{22}^*$ $+ J_{21}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{21}^*)$ $+ J_{12}J_{22}^* - \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{21}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{21}^*)$ $- J_{12}J_{22}^* + \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $+ J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{21}J_{22}^*)$
$\nearrow$	$C - \text{Re}(J_{11}J_{21}^*)$	$D - \text{Re}(J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{11}J_{21}^* - J_{11}J_{22}^*$ $- J_{12}J_{21}^* - J_{12}J_{22}^*$ $+ J_{21}J_{22}^*)$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{11}J_{21}^* - J_{11}J_{22}^*$ $- J_{12}J_{21}^* + J_{12}J_{22}^*$ $+ J_{21}J_{22}^*)$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{21}^*)$ $+ J_{12}J_{22}^* - \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $- J_{11}J_{22}^* + J_{12}J_{21}^*$ $+ J_{21}J_{22}^*)$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{21}^*)$ $- J_{12}J_{22}^* + \frac{1}{2}\text{Im}(J_{11}J_{12}^*)$ $- J_{11}J_{22}^* + J_{12}J_{21}^*$ $+ J_{21}J_{22}^*)$
$\odot$	$C + \text{Im}(J_{11}J_{21}^*)$	$D + \text{Im}(J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{21}J_{22}^* + \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $+ J_{11}J_{22}^* + J_{12}J_{21}^*$ $+ J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{21}J_{22}^* + \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{22}^*)$ $- J_{12}J_{21}^* - \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{21}J_{22}^*)$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{22}^*)$ $- J_{12}J_{21}^* + \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{21}J_{22}^*)$
$\odot$	$C - \text{Im}(J_{11}J_{21}^*)$	$D - \text{Im}(J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{21}J_{22}^* - \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $+ J_{11}J_{22}^* + J_{12}J_{21}^*$ $+ J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{12}^*)$ $- J_{21}J_{22}^* + \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{12}J_{22}^*)$	$\frac{1}{2}(A+B) - \frac{1}{2}\text{Re}(J_{11}J_{22}^*)$ $- J_{12}J_{21}^* - \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $- J_{11}J_{22}^* + J_{12}J_{21}^*$ $+ J_{21}J_{22}^*)$	$\frac{1}{2}(A+B) + \frac{1}{2}\text{Re}(J_{11}J_{22}^*)$ $- J_{12}J_{21}^* + \frac{1}{2}\text{Im}(J_{11}J_{21}^*)$ $- J_{11}J_{22}^* - J_{12}J_{21}^*$ $+ J_{21}J_{22}^*)$

<sup>a</sup>Here we have set  $A = \frac{1}{2}(|J_{11}|^2 + |J_{12}|^2)$ ,  $B = \frac{1}{2}(|J_{21}|^2 + |J_{22}|^2)$ ,  $C = \frac{1}{2}(|J_{11}|^2 + |J_{21}|^2)$ , and  $D = \frac{1}{2}(|J_{12}|^2 + |J_{22}|^2)$ .

[47,48]. We can evaluate the Jones matrix associated with such material medium by utilizing Eq. (5.1.32). The result is

$$\mathbf{J}(0, L) = \begin{bmatrix} \exp\left(\frac{i\pi vL}{c} (\Delta n + i\Delta k)\right) & 0 \\ 0 & \exp\left(-\frac{i\pi vL}{c} (\Delta n + i\Delta k)\right) \end{bmatrix},$$

where  $L$  denotes the sample thickness. I leave to the reader the task of proving that

$$(\leftrightarrow, \leftrightarrow) = \exp\left(-\frac{2\pi L}{c} \Delta k\right) \quad (5.4.3a)$$

$$[(\curvearrowleft, \curvearrowleft) - (\curvearrowright, \curvearrowright)] = \sin\left(\frac{2\pi vL}{c} \Delta n\right) \quad (5.4.3b)$$

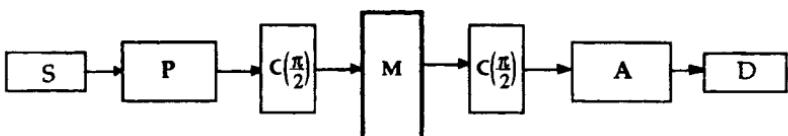
$$[(\curvearrowleft, \curvearrowright) - (\curvearrowright, \curvearrowleft)] = \cos\left(\frac{2\pi vL}{c} \Delta n\right) \quad (5.4.3c)$$

Since the birefringence appears as the argument of a multivalued sinusoidal function, data taken at a single frequency cannot be uniquely inverted. This difficulty may be circumvented through the use of polychromatic illumination since different wavelengths will be extinguished at different values of  $\Delta n$ , or by varying the thickness of the sample.

We have, up to this point, focused only on optical systems that can be characterized by Jones matrices. We now turn to the characterization of optical elements by Muller matrices.

#### 5.4.2. METHODS FOR MEASURING THE MUELLER MATRICES

It is the purpose of this section to investigate how the Mueller matrix of an arbitrary linear optical device can be measured. For completeness and maximum generality, we assume that this optical device can transform and/or depolarize the incident beam, as well as change its direction of propagation. The characteristic Mueller matrix for any optical device or scattering medium contains 16 elements, so it requires 16 equations to solve for these elements. Hence, the Mueller matrix is calculable from 16 intensity measurements. The set of these 16 elements contains all the information obtainable from the optical system at this frequency and, in the case of light scattering, at given scattering angle. In practice, all 16 elements may not be independent. Some are zero, and some are identical to others depending on the symmetry and certain properties of the optical medium. For certain ensembles of perfect particles such as spheres, the matrix elements can be exactly predicted (see Section 4.3). In this section, two types of optical arrangements are described and analyzed



**Figure 5.4.2.** The different components of the experimental setup: the light source  $S$ , a polarizing unit  $P$ , the optical medium characterized by a Mueller matrix  $M$ , the analyzing unit  $A$ , and the radiation detector  $D$ .

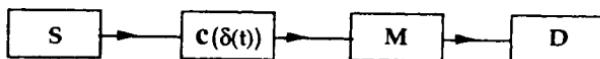
for the purpose of measuring the full Mueller matrix: (1) a transmission experiment and (2) methods that introduce a modulation of the polarization. For complete discussions of these methods, the reader may consult, for example, Refs. [49–63].

#### 5.4.2.1. Transmission Experiment

Figure 5.4.2 shows a conventional scheme of the optical components that is used for these measurements. A light source emits light that is directed through a polarizer and an input optics. Each light beam can be prepared into a pure state of polarization. The light then passes through a quarter-wave linear retarder whose orientation is varied in angular increments. These polarization states interact with the sample and then exit it. The beam leaving the sample is further analyzed by a second quarter-wave retarder and a linear polarizer before reaching an optical detector. A minimum of 16 incident and analyzed polarization states are required to determine the Mueller matrix. A variety of other types of transmission polarimeters is presented in Fuller [51]. Several possible sources of systematic errors are present in this measurement procedure. In practice, small errors in the orientation of the polarizers and in the retardance of the compensators result in large errors in the measured Mueller matrix.

#### 5.4.2.2. Polarization Modulation Method

One of the most employed methods of measurement of the Mueller matrix is by means of photoelastic modulation. This technique offers a number of advantages, which are chronicled in the review of Jasperson and Schnatterly [49]. In brief a photoelastic modulator (PEM) uses the photoelastic effect to produce polarization modulation of light. Acoustical resonance produces an oscillating birefringence at a frequency determined by the physical characteristics of the device. For pioneering studies and applications of PEM, we refer the reader to Refs. [42–53]. Kemp and coworkers have measured polarization components to less than one part per million (<1 ppm) of the background light intensity [52].



**Figure 5.4.3.** Block diagram of the experimental setup, where  $S$  is a quasimonochromatic light source,  $C(\delta(t))$  is a phase modulator,  $M$  the Mueller matrix to be determined, and  $D$  is the detection system.

The instrument, called a *polarization-modulated nephelometer*, uses the polarization modulation technique, with the Bessel function method, developed by Hunt and Huffman [50]. The schematics of the instrument is shown in the block diagram of Fig. 5.4.3. It operates by periodically modulating the incident beam's polarization state by means of a variable phase retarder (i.e., photoelastic modulator) while observing the signals carried by the fundamental and different harmonics of the output light beam. The PEM in the optical setup may be a birefringent medium of appropriate thickness, such as fused-silica quartz in which a compressional standing sound wave is induced to generate a sinusoidally time-varying strain of angular frequency  $w$ . The change in the index of refraction is a linear function of the strain. The phase retardation  $\delta$  between the 1 and 2 components of the light propagating along  $e_3$  is also a sinusoidally varying function,  $\delta(t) = A \sin(wt)$ . Here  $A$  is the amplitude of the retardance of the modulator.

We now derive the equations for the intensity of the light reaching the detector as a function of time.

$$\mathbf{S}_o = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \\ m_{30} & m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos(\delta) & \sin(\delta) \\ 0 & 0 & -\sin(\delta) & \cos(\delta) \end{bmatrix} \mathbf{S}_i \quad (5.4.4)$$

Now we can expand  $\sin(\delta)$  and  $\cos(\delta)$  in Fourier series:

$$\sin(\delta(t)) = \sin(A \sin(wt)) = 2 \sum_{k=0}^{\infty} J_{2k+1}(A) \sin((2k+1)wt) \quad (5.4.5a)$$

$$\cos(\delta(t)) = \cos(A \sin(wt)) = J_0(A) + 2 \sum_{k=1}^{\infty} J_{2k}(A) \cos(2kw) \quad (5.4.5b)$$

where the  $J_k$  terms are the Bessel functions of first kind and  $k$ th order. We further choose  $A = A^* = 2.405$  rad ( $\cong 138^\circ$ ) such that  $J_0(A^*) = 0$ , in order to have a dc-free Fourier expansion in Eq. (5.4.5b). Specific combinations of the Mueller matrix elements are encoded on the different frequencies of the modulated output intensity. If the series occurring in Eqs. (5.4.5a,b) are truncated to second order, we can prove that the gain and output Stokes

parameter  $\langle S_1 \rangle_o$ , as a function of time, are given by

$$\begin{aligned} g &= m_{00} \pm m_{01} \\ \langle S_1 \rangle_o &= m_{10} \pm m_{11} \end{aligned} \quad (5.4.6a)$$

if we set

$$\mathbf{S}_i = \begin{bmatrix} 1 \\ \pm 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} g &= m_{00} \pm 2(J_2(A^*) \cos(2wt)m_{02} - J_1(A^*) \sin(wt)m_{03}) \\ \langle S_1 \rangle_o &= m_{10} \pm 2(J_2(A^*) \cos(2wt)m_{12} - J_1(A^*) \sin(wt)m_{13}) \end{aligned} \quad (5.4.6b)$$

if we set

$$\mathbf{S}_i = \begin{bmatrix} 1 \\ 0 \\ \pm 1 \\ 0 \end{bmatrix}$$

and

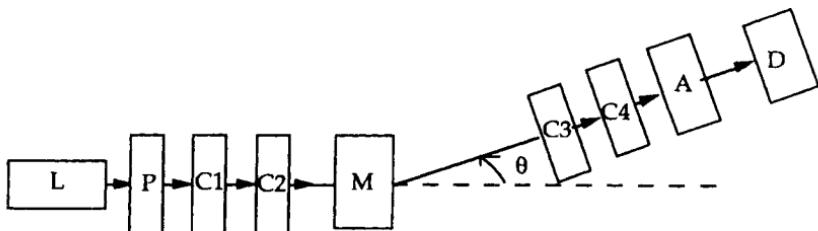
$$\begin{aligned} g &= m_{00} \pm 2(J_2(A^*) \cos(2wt)m_{03} - J_1(A^*) \sin(wt)m_{02}) \\ \langle S_1 \rangle_o &= m_{10} \pm 2(J_2(A^*) \sin(wt)m_{12} - J_1(A^*) \cos(2wt)m_{13}) \end{aligned} \quad (5.4.6c)$$

if we set

$$\mathbf{S}_i = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \pm 1 \end{bmatrix}$$

We remark that the output Stokes parameter  $\langle S_1 \rangle_o$  is composed of a dc component and sinusoidally time-varying components, either at the first harmonics  $w$  or second harmonics  $2w$ . We left to the reader the task of verifying that similar formulas apply for the other Stokes parameters. Using a reference signal, these harmonics can be detected using lock-in amplification. Another point of interest is that no harmonics are present in Eq. (5.4.6a) corresponding to linear horizontal and vertical input Stokes vectors. Thence, by analyzing with a polarizer  $A$ , selected  $m_{kj}$  elements, or specific  $m_{kj}$  combinations can be chosen.

In practice, the method of measuring the Mueller matrix elements by using a single modulator configuration turns out to be rather laborious. For this



**Figure 5.4.4.** Four-modulator photopolarimeter of Thompson, Bottiger, and Fry [53].

reason, a four-modulator photopolarimeter has been developed by Thomson, Bottiger, and Fry, permitting the simultaneous measurement of the Mueller matrix elements with an absolute accuracy of 1–5%. With an experimental arrangement something like Fig. 5.4.4 in mind, we now describe the design of this polarimeter. A beam of quasimonochromatic well-collimated light is incident on the polarimeter. The instrument has two electrooptic modulators (Pockels cell modulators) that modulate the polarization of the incident light beam for different states of polarization at audiofrequencies. Then, the transmitted (or scattered) light passes through a cascade of two electrooptic modulators and a linear polarizer before reaching a photodetector. Within this arrangement, the Mueller matrix elements can be measured as the gain at different audiofrequencies. Following the formalism developed for the single-modulator configuration, I have summarized in Eq. (5.4.7) the linear combinations of frequencies at which each matrix element appears in the expression of the gain.

$$\mathbf{M} = \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ (\text{dc}) & (2w_1) & (w_1 \pm w_2) & (w_1 \pm 2w_2) \\ m_{10} & m_{11} & m_{12} & m_{13} \\ (2w_4) & 2(w_1 \pm w_4) & (w_1 \pm w_2 \pm 2w_4) & (w_1 \pm 2w_2 \pm w_4) \\ m_{20} & m_{21} & m_{22} & m_{23} \\ (w_3 \pm w_4) & (2w_1 \pm w_3 \pm w_4) & (w_1 \pm w_2 \pm w_3 \pm w_4) & (w_1 \pm 2w_2 \pm w_3 \pm w_4) \\ m_{30} & m_{31} & m_{32} & m_{33} \\ (2w_3 \pm w_4) & (2w_1 \pm 2w_3 \pm w_4) & (w_1 \pm w_2 \pm 2w_3 \pm w_4) & (w_1 \pm 2w_2 \pm 2w_3 \pm w_4) \end{bmatrix} \quad (5.4.7)$$

where  $w_j$  ( $j = 1, 2, 3, 4$ ) are the primary frequencies of vibration of the four PEMs.

For example, the frequencies at which  $m_{33}$  appears in the gain are  $w_1 \pm 2w_2 \pm 2w_3 \pm w_4$ . Bear in mind that when  $J_0(A^*) = 0$ , the dc intensity becomes a constant, that is, independent of changes that could be made to the optical system and therefore can be used for normalization of the ac signals by forming the intensity ac/intensity dc ratio. This ratio is now independent of fluctuations in intensity of the light source. Typical accuracy is on the order of

percent. As emphasized and discussed in Ref. [53], the dominant error for this arrangement is systematic and is due mainly to the misalignment of the optical components of the system. Specifically, this causes mixing in the neighboring matrix elements. Other errors arise from imperfect optical components and a partially polarizing detector. The polarization modulation method of measurement of the full Mueller matrix is nondestructive, fast, and easily calibrated. It has found applications in many fields, including suspensions of hydrosols [53].

We have, up to this point, concentrated on nonimaging Mueller matrix polarimeters. In fact, polarization can provide a richer set of descriptive physical constraints than does intensity, for the interpretation of an imaged scene. The *Mueller matrix imaging polarimeter* (MMIP) measures the spatially dependent polarization properties of optical systems. The imaging polarimeter is composed of a standard polarimeter with the addition of a charge-coupled device (CCD) detector to measure an array of Mueller matrices associated with multiple ray paths through an optical system; thus, it measures the polarization of scattered (or transmitted) light as Mueller matrix images. Several MMIP have been described in the literature for various applications, including earth resource investigation and remote sensing [64]. Another important aspect for interpreting objects in a scene is identification of material classification. The capability of determining whether parts of an object are metal or dielectric can be very useful to object recognition and material inspection in manufacturing [65].

As we have seen, a problem of practical importance in polarimetry is to determine the Mueller matrix elements from selected measurements of Stokes vectors, specifically, calibration states. In principle, a minimum of 16 incident and analyzed states of polarization are needed to determine the Mueller matrix, but in practice the measurement procedure often operates with a larger set of incident polarization states (e.g., 64), overdetermining the Mueller matrix so one seeks the set of independent matrix elements  $m_{ij}$  that best fits the measurements in a least-squares sense. A limited research work on this subject has been reported in the scientific and engineering literatures [66–69].

There has been a great deal of practical and theoretical interest in polarization optics, and one expects continued experimental and theoretical efforts in this field. The applications of these works cover an enormous range, and we have been able here to look at a limited corner of different areas. The limitation of available space requires that many interesting and important topics be excluded from our discussion, such as optical systems of encoding and decoding information using polarization of light. From our personal perspective, we anticipate many yet-unthought-of developments in these and other areas. Our hope is that the interest in the basic physics of polarization will produce advances that cannot only lead to improved applications for technological processes but also deepen our understanding of the general principles of optics.

## FURTHER READING

1. See the feature issue on wave propagation and scattering in random media, *J. Opt. Soc. Am.* **A2**, 1985, 206 and references cited therein.
2. J. M. Bennett and H. E. Bennett, "Polarization," in *Handbook of Optics*, W. G. Driscoll and W. Vaughan, eds., McGraw-Hill, New York, 1979.
3. R. M. A. Azzam and N. M. Bashara, *Ellipsometry and Polarized Light*, North-Holland, Amsterdam, 1977.
4. A. Yariv and P. Yeh, *Optical Waves in Crystals*, Wiley, New York, 1984.
5. F. R. Stauffer, "Optical polarization," in *Methods of Experimental Physics*, D. Malacara, ed., Academic Press, New York, 1988, Vol. 26.
6. G. N. Ramachandran and S. Ramaseshan, "Crystal optics," in *Handbuch der Physik*, S. Flügge, ed., Springer-Verlag, Berlin, 1961, Vol. XXV.
7. I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series and Products*, Academic Press, New York, 1965.
8. R. C. Jones, *J. Opt. Soc. Am.* **48**, 1948, 671.
9. R. M. A. Azzam, *J. Opt. Soc. Am.* **68**, 1978, 1756.
10. R. M. A. Azzam and N. B. Bashara, *J. Opt. Soc. Am.* **62**, 1972, 1252. See also B. E. Merrill, R. M. A. Azzam, and N. M. Bashara, *J. Opt. Soc. Am.* **64**, 1974, 731.
11. R. J. Vernon and B. D. Huggins, *J. Opt. Soc. Am.* **70**, 1980, 1364.
12. C. Brosseau, *Opt. Lett.* **20**, 1995, 1221.
13. P. G. de Gennes and J. Prost, *The Physics of Liquid Crystals*, 2nd ed., Clarendon, Oxford, 1993.
14. C. Booth and P. Raynes, *Physics World* **1997**, 33. See also S. Chandrasekhar, *Liquid Crystals*, 2nd ed., Cambridge Univ. Press, Cambridge, UK, 1992.
15. See, for example, W. H. Press, B. P. Flannery, S. A. Teukolshy, and W. T. Vetterling, *Numerical Recipes*, Cambridge Univ. Press, Cambridge, UK, 1986.
16. D. Marcuse, *Theory of Optical Waveguides*, Academic Press, New York, 1974.
17. S. C. Rashleigh, *IEEE J. Lightwave Technol.* **LT-1**, 1983, 312 (fiber optics).
18. I. Kaminow, *IEEE J. Quantum Elect.* **QE-17**, 1981, 15.
19. T. Okoshi, *Optical Fibers*, Academic Press, New York, 1982.
20. T. Eftimov and T. Kortenski, *J. Mod. Opt.* **36**, 1989, 287.
21. W. K. Burns, R. P. Moeller, and Chin Lin Chen, *J. Lightwave Technol.* **LT-1**, 1983, 44.
22. J. M. Bennett, "Polarizers," in *Handbook of Optics*, 2nd ed., M. Bass, ed., McGraw-Hill, New York, 1994.

23. W. A. Shurcliff, *Polarized Light*, Harvard Univ. Press, Cambridge, MA, 1962.
24. W. J. Silfast, *Laser Fundamentals*, Cambridge Univ. Press, New York, 1996.
25. S. M. MacNeille, "Beam splitter," U.S. Patent 2,403,731 July 9, 1946.
26. L. Songer, *Opt. Spectra* **12**, 1978, 45.
27. D. L. Steinmetz, W. G. Phillips, M. Wirick, and F. F. Forbes, *Appl. Opt.* **6**, 1967, 1001.
28. D. Blanc, P. H. Lissberger, and A. Roy, *Thin Solid Films* **57**, 1979, 191.
29. G. R. Bird and M. Parrish, Jr., *J. Opt. Soc. Am.* **50**, 1960, 886.
30. E. H. Land, *J. Opt. Soc. Am.* **41**, 1951, 957. See also E. H. Land and C. D. West, *Colloid Chem.* **6**, 1946, 160 for details about the discovery of dichroic polarizers.
31. I. G. Wood and A. M. Glazer, *J. Appl. Cryst.* **13**, 1980, 217.
32. A. M. Title, *Appl. Opt.* **14**, 1975, 229. See also S. Pancharatnam, *Proc. Indian Acad. Sci.* **A41**, 1955, 137.
33. P. S. Hauge, "Techniques of measurement of the polarization-altering properties of linear optical systems," in *Optical Polarimetry: Instrumentation and Applications*, R. M. A. Azzam and D. L. Coseen, eds., *Proc. Soc. Photo-Opt. Instrum. Eng.* **112**, 1977, 2. See also a survey of methods for the complete determination of a state of polarization: P. S. Hauge, W. L. Hyde, and R. M. A. Azzam, eds., *Proc. Soc. Photo-Opt. Instrum. Eng.* **88**, 1976, 3.
34. A. M. Shutov, *Sov. J. Opt. Technol.* **52**, 1985, 682.
35. H. G. Berry, G. Gabriele, and A. E. Livingston, *Appl. Opt.* **16**, 1977, 3200.
36. R. M. A. Azzam, *Opt. Acta*, **29**, 1982, 629.
37. A. F. Fercher and P. F. Steeger, *Opt. Acta*, **28**, 1981, 443. See also A. F. Fercher and P. F. Steeger, *Opt. Acta* **39**, 1984, 1395; P. F. Steeger, T. Asakura, K. Zocha, and A. F. Fercher, *J. Opt. Soc. Am. A1*, 1984, 677 and P. F. Steeger, *Opt. Lett.* **8**, 1983, 528.
38. I. Freund, *Opt. Commun.* **81**, 1991, 251. See also I. Freund, M. Rosenbluh, and S. Feng, *Phys. Rev. Lett.* **61**, 1988, 2328 and I. Freund, M. Kaveh, R. Berkovits, and M. Rosenbluh, *Phys. Rev. B* **42**, 1990, 2613.
39. C. V. Kent and J. Lawson, *J. Opt. Soc. Am.* **27**, 1937, 117.
40. E. Collett, *Opt. Commun.* **52**, 1984, 77.
41. D. S. Kliger, J. W. Lewis, and C. E. Randall, *Polarized Light in Optics and Spectroscopy*, Academic Press, New York, 1990.
42. R. M. A. Azzam, *Optica Acta*, **29**, 1982, 685; ibid. **32**, 1985, 767. See also R. M. A. Azzam, *Opt. Lett.* **10**, 1985, 309; R. M. A. Azzam, I. M. Elminyawi, and A. M. El-Saba, *J. Opt. Soc. Am. A5*, 1988, 688; R. M. A. Azzam and A. G. Lopez, *J. Opt. Soc. Am. A6*, 1989, 1513.
43. S. Krishnan, *J. Opt. Soc. Am. A9*, 1992, 1615. See also S. Krishnan and P. C. Nordine, *Appl. Opt.* **33**, 1994, 4184.
44. R. C. Jones, *J. Opt. Soc. Am.* **38**, 1948, 671.
45. R. E. Raab, *Optica Acta*, **29**, 1982, 1243.
46. K. Brudzewski, *J. Mod. Opt.* **38**, 1991, 889.
47. C. Brosseau, *C. R. Acad. Sci. (Paris)* **302**, Série II, 1986, 1047. See also C. Brosseau, *Optik* **72**, 1985, 38.

48. H. Itoh, Y. Ishijima, K. Takeshigo, S. Kanba, and K. Sasaki, *Appl. Opt.* **26**, 1987, 17.
49. S. N. Jasperson and S. E. Schnatterly, *Rev. Sci. Instrum.* **40**, 1969, 761.
50. A. J. Hunt and D. R. Huffman, *Rev. Sci. Instrum.* **44**, 1973, 1753. See also R. J. Perry, A. J. Hunt, and D. R. Huffman, *Appl. Opt.* **17**, 1978, 2700.
51. G. G. Fuller, *Optical Rheometry of Complex Fluids*, Oxford Univ. Press, New York, 1995.
52. J. C. Kemp, *J. Opt. Soc. Am.* **59**, 1969, 950. See also J. C. Kemp, G. D. Henson, C. T. Steiner, and E. R. Powell, *Nature* **326**, 1987, 270.
53. R. C. Thompson, J. R. Bottiger, and E. S. Fry, *Appl. Opt.* **19**, 1980, 1323. See also R. C. Thompson and E. S. Fry, *J. Opt. Soc. Am.* **64**, 1974, 1399 and R. C. Thompson, *J. Opt. Soc. Am.* **66**, 1976, 1134.
54. R. G. Johnston, S. B. Singham, and G. C. Salzman, *Appl. Opt.* **25**, 1986, 3566. See also R. G. Johnston, S. B. Singham, and G. C. Salzman, *Comments Mol. Cell. Biophys* **5**, 1988, 171.
55. G. F. Beardsley, *J. Opt. Soc. Am.* **58**, 1968, 52.
56. A. C. Holland and G. Gagne, *Appl. Opt.* **9**, 1970, 1113.
57. D. E. Aspnes, *Opt. Commun.* **8**, 1973, 222.
58. P. S. Hauge, *J. Opt. Soc. Am.* **68**, 1978, 1519.
59. R. M. A. Azzam, *Opt. Lett.* **5**, 1978, 148.
60. S. Asano and M. Sato, *Appl. Opt.* **19**, 1980, 962.
61. W. S. Bickel and W. M. Bailey, *Am. J. Phys.* **53**, 1985, 468.
62. R. Anderson, *Appl. Opt.* **31**, 1992, 11.
63. W. M. McClain, Wen-Haw Jeng, Biswajit Pati, Yaoming Shi, and Duan Tian, *Appl. Opt.* **33**, 1230, 1994.
64. W. G. Egan, J. Grusauskas, and H. B. Hallock, *Appl. Opt.* **7**, 1968, 1529.
65. L. B. Wolff, *J. Opt. Soc. Am.* **11**, 1994, 2935.
66. See the feature issue "Polarization Analysis and Measurement," *Opt. Eng.* **34**, 1995, 1543 and references cited therein.
67. D. G. M. Anderson and R. Barakat, *J. Opt. Soc. Am.* **A11**, 1994, 2305.
68. M. A. F. Thiel, *J. Opt. Soc. Am.* **66**, 1976, 65.
69. E. Landi Degl'Innocenti and J. C. del Toro Iniesta, *J. Opt. Soc. Am.* **A15**, 1998, 533.

---

---

## **APPENDIXES**

# **Analogy of Second-Order Coherence Properties of Blackbody Radiation with Theory of Homogeneous and Isotropic Turbulence of an Incompressible Fluid**

In Section 3.1, second-order tensors that characterize the correlations existing between the electromagnetic field vectors at any two points in the field, at any two instants of time, were introduced. The equations of propagation of these second-order correlation tensors, derived from Maxwell's equations, were also given. Here an analogy with the theory of isotropic turbulence of an incompressible fluid is discussed. The original derivation for this remarkable analogy was provided by Bourret [1].

By formulating the problem as in Section 3.1.5.2, the turbulent motion is determined statistically by the complete system of joint PDF of the vector velocity  $\mathbf{v}(\mathbf{R}, t)$  at any  $N$  points of space-time. However, a main difference is that: turbulence is specified at a definite instant of time, and this requires a knowledge of the complete system of joint PDF of the values of the velocity  $\mathbf{v}$  at any  $N$  points of space at the appropriate value of  $t$ . Turbulence is homogeneous if the flow has identical statistical properties at all points in space; in other words, the PDFs depend only on relative positions of the  $N$  points in space, and are independent of the location in space of the configuration of  $N$  points. In what follows we consider only the case  $N = 2$ . It is isotropic if, in addition, the statistical properties of the flow are the same in all directions. Then the mean velocity  $\langle \mathbf{v}(\mathbf{R}, t) \rangle$  is space- and time-independent and may be chosen zero if one studies such a turbulence in a frame moving with the mean flow. In like fashion as we are concerned in polarization optics with the second-order electric field correlation tensor  $\Phi$ , we define the second-order velocity correlation tensor  $F$ , which is the working tool of the subject. Both tensors obey the same (Helmholtz) equation of propagation. In order to point out, with maximum clarity, the analogies between the formalisms that serve to describe the coherence properties of blackbody radiation and the kinematics of spatially homogeneous and isotropic turbulence, we have listed, in Table A.1, general results concerning the electric field and velocity correlation tensors that are of special physical importance to our purpose. The

**TABLE A.1. Comparison of Basic Equations for Electric Field and Velocity Vectors and Their Respective Correlation Tensors<sup>a</sup>**

Electric Field	Velocity Field
$\operatorname{div} \mathbf{E} = 0$	$\operatorname{div} \mathbf{v} = 0$
$\langle \mathbf{E}(\mathbf{R}, t) \rangle = 0$	$\langle \mathbf{v}(\mathbf{R}, t) \rangle = 0$
$\Phi_{jk}(\mathbf{R}, \tau) = \langle E_j(\mathbf{R}, t + \tau) E_k^*(\mathbf{R}, t) \rangle$	$F_{jk}(\mathbf{R}, t) = \langle v_j(\mathbf{x} + \mathbf{R}, t) v_k(\mathbf{x}, t) \rangle$
$\Phi_{jk}(\mathbf{R}, \tau) = \Phi_{kj}^*(\mathbf{R}, -\tau)$	$F_{jk}(\mathbf{R}, t) = F_{kj}(-\mathbf{R}, t)$
$ \Phi_{jk}(\mathbf{R}, \tau) ^2 \leq \Phi_{jj}(\mathbf{R}, 0) \Phi_{kk}(\mathbf{R}, 0)$	$ F_{jk}(\mathbf{R}, t) ^2 \leq F_{jj}(0, t) F_{kk}(0, t)$
$\nabla^2 \Phi_{jk}(\mathbf{R}, \tau) = \frac{1}{c^2} \frac{\partial^2 \Phi_{jk}(\mathbf{R}, \tau)}{\partial \tau^2}$	$\nabla^2 F_{jk}(\mathbf{R}, t) = \frac{1}{c_s^2} \frac{\partial^2 F_{jk}(\mathbf{R}, t)}{\partial t^2}$

<sup>a</sup>The term  $c_s$  is the sound velocity in the fluid.

reader who is interested in a more complete discussion of the kinematics of homogeneous turbulence should consult Ref. [2].

The solution of the equation of propagation for the correlation tensor  $\Phi$ , or  $\mathbf{F}$ , is a purely mathematical problem. The main point discussed by Bourret considers symmetry conditions on  $\Phi$  for blackbody radiation. See Section 3.1.5.7 for the details of the basic assumptions. From the nature of stationarity and isotropic fluctuations, it follows that  $\Phi_{jk}(\mathbf{R}, \tau)$  must be even functions of the arguments  $\mathbf{R}$  and  $\tau$ . Hence we may write

$$\Phi_{jk}(\mathbf{R}, \tau) = \int \exp(i\mathbf{q} \cdot \mathbf{R}) \cos(qc\tau) f_{jk}(\mathbf{q}) d\mathbf{q} \quad (\text{A.1})$$

with  $f_{jk}(\mathbf{q}) = f_{jk}(-\mathbf{q})$ . In the case of turbulence of an incompressible fluid ( $\operatorname{div} \mathbf{v} = 0$ ) that has spherical symmetry, the general form of  $f_{jk}(\mathbf{q})$  is

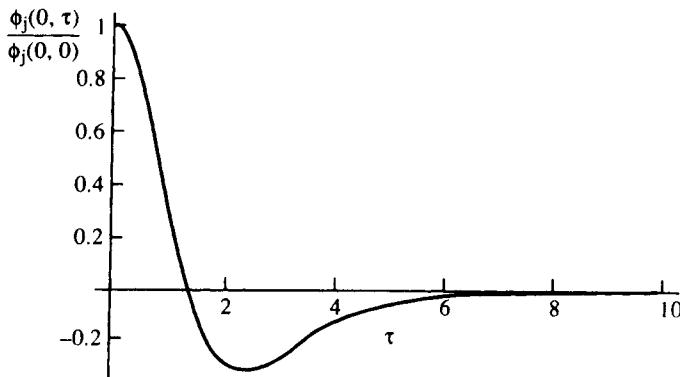
$$f_{jk}(\mathbf{q}) = A(q)(q^2 \delta_{jk} - q_j q_k) \quad (\text{A.2})$$

where the scalar function can be written

$$qA(q) = \frac{2\hbar c}{3\pi} \frac{1}{\exp(a) - 1} \quad (\text{A.3})$$

and the length  $a$  is given by  $a \equiv (\hbar c/kT)$  [1]. By substituting Eq. (A.3) in Eq. (A.1), Bourret was able to obtain the following expression for the electric field correlation tensor:

$$\Phi_{jk}(0, \tau) = -\left(\frac{4}{3}\right)^2 \frac{\hbar}{c^2} \delta_{jk} \frac{\partial^3}{\partial t^3} \int_0^\infty \frac{\sin(qc\tau) dq}{\exp(aq) - 1} = -\frac{8}{9} \hbar c \left(\frac{\pi}{a}\right)^4 \delta_{jk} \mathbf{L}''' \left(\frac{\pi c}{a} \tau\right) \quad (\text{A.4})$$



**Fig. A.1.** Normalized electric field correlation tensor of blackbody radiation as a function of the dimensionless time  $\pi k T \tau / \hbar$ .

where  $L(x) \equiv \cotanh(x) - (1/x)$  is the usual Langevin function [1]. This behavior is illustrated in Fig. A1.

## REFERENCES

1. R. C. Bourret, *Nuovo Cimento* **18**, 1960, 347.
2. There are numerous textbooks devoted to the presentation of the statistical theory of turbulence. See, for instance, L. D. Landau and E. M. Lifschitz, *Fluid Dynamics*, Addison-Wesley, New York, 1959 and G. K. Batchelor, *The Theory of Homogeneous Turbulence*, Cambridge Univ. Press, Cambridge, 1967.

## APPENDIX B

# Derivation of Eq. (3.4.5) Using the Spectral Decomposition Theorem

Here, we introduce the basic theorem of spectral decomposition, for a positive definite matrix, to evaluate the function  $S(P)$ . The interested reader may refer to Graybill [1] and Gantmacher [2] for the details of the different proofs. Let  $\mathbf{X}$  be an  $a \times a$  X matrix. It is said to be definite positive if and only the  $i$ -th leading principal minors of  $\mathbf{X}$  are positive. Now, given any positive definite matrix  $\mathbf{X}$ , there exists  $\mathbf{A}_i, i \in \langle 1, k \rangle$  symmetric idempotent matrices (i.e.,  $\mathbf{A}_i \mathbf{A}_j = \mathbf{A}_i^2 \delta_{ij}$ ) and  $a_i$  positive scalars such that:

$$\mathbf{X} = a_0 \mathbf{O}_0^{(a)} + \sum_{i=1}^k a_i \mathbf{A}_i \quad (\text{B.1})$$

where  $\mathbf{O}_0^{(a)}$  is the notation of the  $a \times a$  unit matrix (e.g.,  $\mathbf{O}_0^{(2)} = \sigma_0$ ). It can also be shown that  $\lambda_0 = a_0$  and  $\lambda_i = a_0 + a_i, i \in \langle 1, k \rangle$  are the distinct eigenvalues of  $\mathbf{X}$ . Next, we want to have an operational formula for the entropy of  $\mathbf{X}$ . Taking the logarithm of both sides (when  $a_0 \neq 0$ ), we have

$$\ln(\mathbf{X}) = \mathbf{O}_0^{(a)} \ln(a_0) + \ln\left(1 + \sum_{i=1}^k \left(\frac{a_i}{a_0}\right) \mathbf{A}_i\right) \quad (\text{B.2})$$

Now, examining in more details the second term of the right-hand side of Eq. (B.2), we find that

$$\ln\left(\mathbf{O}_0^{(a)} + \sum_{i=1}^k \left(\frac{a_i}{a_0}\right) \mathbf{A}_i\right) = \sum_{i=1}^k \ln\left(1 + \left(\frac{a_i}{a_0}\right)\right) \mathbf{A}_i \quad (\text{B.3})$$

where the convergence of the logarithmic series is guaranteed since it is easily seen from the Lagrange–Sylvester theorem that every  $\mathbf{X}$  eigenvalue has its modulus inferior to 1. After simple manipulations, we obtain an explicit evaluation of  $\text{tr}(\mathbf{X} \ln(\mathbf{X}))$  as

$$\text{tr}(\mathbf{X} \ln(\mathbf{X})) = a_0 \ln(a_0) \text{tr}(\mathbf{O}_0^{(a)}) + \sum_{i=1}^k ((a_0 + a_i) \ln(a_0 + a_i) - a_0 \ln(a_0)) \text{tr}(\mathbf{A}_i) \quad (\text{B.4})$$

The usefulness of this expression lies in the fact that the right-hand side of (B.4) can be evaluated in terms of the eigenvalues of  $\mathbf{X}$  and by the traces of the  $\mathbf{A}_i$  projectors.

As illustrative examples, we will set  $\mathbf{X} = \mathbf{D}_2$  or  $\mathbf{D}_{4g}$ . We obtain respectively:

$$\text{If } \mathbf{X} = \mathbf{D}_2, \text{tr}(\mathbf{O}_0^{(2)}) = 2, \quad k = 1, \quad \text{we have} \quad a_0 = \frac{(1 - P)}{2}, \\ a_1 = P, \quad \text{tr}(\mathbf{A}_1) = 1 \quad (\text{B.5})$$

and Eq. (B.4) becomes

$$S_2 = -\text{tr}(\mathbf{D}_2 \ln(\mathbf{D}_2)) = -\ln\left(\frac{1}{2}(1 - P)^{(1-P)/2}(1 + P)^{(1+P)/2}\right) = -\ln(s(P)) \quad (\text{B.6})$$

If

$$\mathbf{X} = \mathbf{D}_{4g}, \text{tr}(\mathbf{O}_0^{(4)}) = 4, \quad k = 2,$$

we have

$$a_0 = \frac{1 - P^2}{4}, \quad a_1 = \frac{P(P + 1)}{2}, \\ a_2 = \frac{P(P - 1)}{2}, \quad \text{tr}(\mathbf{A}_1) = \text{tr}(\mathbf{A}_2) = 1 \quad (\text{B.7})$$

and Eq. (B4) becomes

$$S_4 = -\text{tr}(\mathbf{D}_{4g} \ln(\mathbf{D}_{4g})) = -\ln\left(\frac{1}{4}(1 - P)^{1-P}(1 + P)^{1+P}\right) = -\ln(s(P))^2 \\ = -2\ln(s(P)) \quad (\text{B.8})$$

In both cases,  $\text{tr}(\mathbf{D}_a \ln(\mathbf{D}_a))$  is expressed as a continuous decreasing function of a single parameter: the degree of polarization  $P$ .

## REFERENCES

1. F. A. Graybill, *Matrices with Applications in Statistics*, Wadsworth, Belmont, 1983.
2. F. R. Gantmacher, *The Theory of Matrices*, Chelsea Publishing, New York, 1959.

## APPENDIX C

# Degree of Polarization of an Incoherent Mixture of Partially Polarized Light Beams

In the subsequent calculations presented in this appendix, we are interested in finding bounds for the degree of polarization of an incoherent mixture of partially polarized light beams [1–4]. Experimentally this situation is often encountered; as light propagates through optical systems, its polarization state may unintentionally vary, reflecting coatings, for example.

Our derivation will be based on the concavity property of entropy (i.e., if many beams are fitted together, one loses the information that indicates which ensemble a special beam stems from, and therefore entropy increases) and of the linear superposition of polarization states for an incoherent mixture. It follows from Eq. (3.1.159) that the density matrix of the mixture takes the form

$$\mathbf{D} = \sum_{k=1}^N \lambda_k \mathbf{D}_k \quad (\text{C.1})$$

where  $\mathbf{D}_k$  represent the density matrix for the  $k$ th light beam and  $\lambda_k$  are the intensity weights [ $\lambda_k = \text{tr}(\Phi_k)/\text{tr}(\Phi)$  and  $\sum_{k=1}^N \lambda_k = 1, k \in \langle 1, N \rangle$ ] of the different components of the mixture. Using the concavity and superposition properties, one obtains the expression [5]

$$\sum_{k=1}^N \lambda_k S(\mathbf{D}_k) \leq S(\mathbf{D}) \leq \sum_{k=1}^N \lambda_k S(\mathbf{D}_k) - \sum_{k=1}^N \lambda_k \ln(\lambda_k) \quad (\text{C.2})$$

The term  $\sum_{k=1}^N \lambda_k \ln(\lambda_k)$  occurring on the right-hand side of inequality (C.2) represents the mixing entropy contribution. It is easily verified, if we use Eq. (3.4.5), that

$$\prod_{k=1}^N (\lambda_k)^{\lambda_k} \prod_{k=1}^N s(P_k)^{\lambda_k} \leq s(P) \leq \prod_{k=1}^N s(P_k)^{\lambda_k} \quad (\text{C.3})$$

where the function  $s(x)$  is as defined by Eq. (3.4.6). Because  $\prod_{k=1}^N s(P_k)^{\lambda_k}$  takes

a value between  $\frac{1}{2}$  and 1, there exists a unique value noted  $P_\pi$  verifying the following:

$$P_\pi = s^{-1} \left( \prod_{k=1}^N (s(P_k))^{\lambda_k} \right) \quad (\text{C.4})$$

Hence the result

$$s(P_\pi) \prod_{k=1}^N (\lambda_k)^{\lambda_k} \leq s(P) \leq s(P_\pi) \quad (\text{C.5})$$

Two cases may occur. First, if  $\frac{1}{2} \leq s(P_\pi) \prod_{k=1}^N (\lambda_k)^{\lambda_k} \leq 1$ , then there exists a unique value  $P^*$  given by

$$P^* = s^{-1} \left( s(P_\pi) \prod_{k=1}^N (\lambda_k)^{\lambda_k} \right) \quad (\text{C.6})$$

with  $P^* \leq P \leq P_\pi$ . On the other hand, if

$$0 < s(P_\pi) \prod_{k=1}^N (\lambda_k)^{\lambda_k} < \frac{1}{2} \quad \text{we have} \quad 0 \leq P \leq P_\pi$$

Because the function  $x^\alpha$  with  $0 < \alpha \leq 1$  is bijective strictly increasing, we readily find that

$$P \leq P_\pi \leq P_{\max} \quad (\text{C.7})$$

where  $P_{\max}$  is defined as the extremum of the degrees of polarization of the different beams. The theorem we have just established may also be expressed by saying that when  $N$  quasimonochromatic light beams (of same mean frequency) propagating along the same direction are combined, the degree of polarization of the resultant beam can never be greater than that of the constituent beams. Note that this result is more stringent than the one proposed by Bohren (i.e.  $P \leq P_{\max}$ ) [3]. For concreteness we can treat the case of mixing of two pure states (i.e.,  $P_1 = P_2 = 1$ ) of equal intensity ( $\lambda_1 = \lambda_2 = \frac{1}{2}$ ). It follows from above that  $P_\pi = 1$ , hence the result  $0 \leq P \leq 1$ . Consequently the resultant wave is a mixed state (i.e., a finite polarization temperature), except when the two states are orthogonally polarized for which the mixture leads to unpolarized light or when the two density matrices are identical for which  $P$  remains equal to 1.

**REFERENCES**

1. G. G. Stokes, in *Mathematical and Physical Papers*, Johnson Reprint, New York, 1966, Vol. 3. See also G. G. Stokes, *Trans. Camb. Phil. Soc.* **9**, 1852, 399.
2. G. N. Ramachandran and S. Ramaseshan, "Crystals optics," in *Handbuch der Physik*, S. Flügge, ed., Springer-Verlag, Berlin, 1961, Vol. XXV.
3. C. F. Bohren, *Appl. Opt.* **26**, 1987, 606.
4. C. Brosseau, *C. R. Acad. Sci. Paris* **310**, Série II, 1990, 181.
5. R. Werhl, *Rev. Mod. Phys.* **50**, 1978, 221.

## APPENDIX D

# Set of Generalized Stokes Parameters

In the standard representation of  $SU(3)$  the Gell–Mann matrices are

$$\mathbf{O}_0^{(3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{O}_1^{(3)} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{O}_2^{(3)} = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (\text{D.1a})$$

$$\mathbf{O}_3^{(3)} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{O}_4^{(3)} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}, \quad \mathbf{O}_5^{(3)} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad (\text{D.1b})$$

$$\mathbf{O}_6^{(3)} = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}, \quad \mathbf{O}_7^{(3)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{O}_8^{(3)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} \quad (\text{D.1c})$$

For  $N = 3$ , we may construct a typical set of normalized Stokes parameters by analogy with the case of plane waves  $N = 2$ . They can be written as follows:

$$\Theta_0^{(3)} = 1, \quad \Theta_1^{(3)} = \frac{3}{2} \frac{\langle E_1 E_2^* \rangle + \langle E_1^* E_2 \rangle}{\sum_{j=1}^3 \langle |E_j|^2 \rangle}, \quad \Theta_2^{(3)} = \frac{3}{2} i \left( \frac{\langle E_1 E_2^* \rangle - \langle E_1^* E_2 \rangle}{\sum_{j=1}^3 \langle |E_j|^2 \rangle} \right) \quad (\text{D.2a})$$

$$\Theta_3^{(3)} = \frac{3}{2} \frac{\langle |E_1|^2 \rangle - \langle |E_2|^2 \rangle}{\sum_{j=1}^3 \langle |E_j|^2 \rangle}, \quad \Theta_4^{(3)} = \frac{\sqrt{3}}{2} \frac{\langle |E_1|^2 \rangle + \langle |E_2|^2 \rangle - 2\langle |E_3|^2 \rangle}{\sum_{j=1}^3 \langle |E_j|^2 \rangle},$$

$$\Theta_5^{(3)} = \frac{3}{2} \frac{\langle E_1 E_3^* \rangle + \langle E_1^* E_3 \rangle}{\sum_{j=1}^3 \langle |E_j|^2 \rangle} \quad (\text{D.2b})$$

$$\Theta_6^{(3)} = \frac{3}{2} i \left( \frac{\langle E_1 E_3^* \rangle - \langle E_1^* E_3 \rangle}{\sum_{j=1}^3 \langle |E_j|^2 \rangle} \right), \quad \Theta_7^{(3)} = \frac{3}{2} \frac{\langle E_2 E_3^* \rangle + \langle E_2^* E_3 \rangle}{\sum_{j=1}^3 \langle |E_j|^2 \rangle}$$

$$\Theta_8^{(3)} = \frac{3}{2} i \left( \frac{\langle E_2 E_3^* \rangle - \langle E_2^* E_3 \rangle}{\sum_{j=1}^3 \langle |E_j|^2 \rangle} \right) \quad (\text{D.2c})$$

---

## APPENDIX E

---

# Short Historical Account of the Concept of Entropy of Radiation Fields

The purpose of this short historical account is to review the historical basis of the development of the concept of the entropy of a radiation field. The discussion about the relation between thermodynamics and optics is an old and an interesting one. Various aspects of this story have been discussed by Barakat [1–2] and Brosseau [3], to mention only a few. Studies of the entropy of radiation have spanned a long period of time, dating back to the pioneering researches of Wien (1894). Since then, other notable names in physics, such as Planck, von Laue, and many others, have made substantial contributions to this field. The story has culminated in the development of the quantum theory of the electromagnetic radiation in thermal equilibrium.

The foundation of the modern theory of entropy of radiation was laid by Max Planck [4]. Planck's revolutionary idea to describe fundamental physics through the concept of photons not only motivated the first studies on quantum properties of radiation, but also stimulated the interest in the coherence properties of light. Let us begin our excursion by reviewing a few facts about radiation theory. The total radiation energy for linearly polarized light passing through a surface element  $da$  in time  $dt$  is written as [4]

$$da dt \int K_v \cos(\theta) d\Omega dv \quad (\text{E.1})$$

where  $K_v$  is the specific radiation intensity of frequency  $v$  that passes through the surface in the solid angle  $d\Omega$  and  $\theta$  is the angle between the direction of propagation and the normal to the surface. Similarly, Planck wrote down an equation for the radiation entropy as

$$da dt \int L_v \cos(\theta) d\Omega dv \quad (\text{E.2})$$

where  $L_v$  is the specific radiation entropy of frequency  $v$  that passes through the surface in the solid angle  $d\Omega$ . Planck was also able to derive a functional relation between  $K_v$  and  $L_v$  in the form

$$\frac{L_v}{v^2} \equiv f \left( \frac{K_v}{v^3} \right) \quad (\text{E.3})$$

However, no explicit expression of the function  $f$  was given. This was left to his student von Laue, who showed that [5]

$$\frac{L_v}{v^2} = \frac{k}{c^2} [(1 + x_v) \ln(1 + x_v) - x_v \ln(x_v)] \quad (\text{E.4})$$

where  $x_v \equiv c^2/h(K_v/v^3) = 1/[\exp(hv/kT) - 1]$  which is the Planck radiation law for blackbody radiation in an enclosure at equilibrium temperature  $T$ . Here  $k$  is the Boltzmann constant,  $h$  is the Planck constant, and  $c$  is the velocity of light in vacuo. Equation (E.4) was rederived more recently by a number of authors, including Rosen and Ore [6]. The radiation temperature for a completely polarized light beam of frequency  $v$  is defined as

$$\frac{1}{T_v} \equiv \frac{\partial L_v}{\partial K_v} \quad (\text{E.5})$$

Since that time, many researchers have worked on various aspects of this problem. Clark Jones deserves special credit for his contribution to our understanding of the tradeoff between entropy, light polarization, and irreversibility [7]. Recently Gudkov was able, using Planck's spectral law, to calculate the total flux of the spectral entropy of an unpolarized light across the surface of a scattering object [8]. A similar approach, but generalized to a partially polarized radiation field, was put forward by Callies, who derived a formula for the local entropy production rate by a single scattering using the radiative transfer theory [9].

## REFERENCES

1. R. Barakat, *Opt. Acta* **30**, 1983, 1171.
2. R. Barakat and C. Brosseau, *J. Opt. Soc. Am.* **10**, 1993, 529.
3. C. Brosseau, *C. R. Acad. Sci. (Paris)* **B310**, 1990, 181. See also C. Brosseau, *Optik* **88**, 1991, 109.
4. M. Planck, *Vorlesungen über die Theorie der Warmestrahlung*, 5th ed., Barth, Leipzig, 1923.
5. M. von Laue, *Ann. Phys.* **28**, 1907, 1.
6. P. Rosen, *Phys. Rev.* **96**, 1954, 555. See also A. Ore, *Phys. Rev.* **98**, 1955, 887.
7. R. C. Jones, *J. Opt. Soc. Am.* **43**, 1953, 138.
8. N. D. Gudkov, *Opt. Spectrosc.* **68**, 1990, 130.
9. H. Callies, *Beitr. Phys. Atmosph.* **62**, 1989, 212.

## APPENDIX F

# Singular Value Decomposition

The purpose of this appendix is to introduce the singular value decomposition (SVD) theorem [1]. The SVD theorem is used in a wide variety of situations in physics. Its appearance in optics, including image compression and processing, is well known [2, 3]. In SVD, a general matrix is reduced to a diagonal form by premultiplying and postmultiplying it by unitary matrices. The SVD of a complex-valued matrix  $\mathbf{A}$  of size  $N \times N$  is given by the product of three matrices

$$\mathbf{A} = \mathbf{U}\mathbf{B}\mathbf{V}^+ \quad (\text{F.1})$$

where  $\mathbf{U}$  and  $\mathbf{V}$  are unitary matrices of size  $N \times N$ . The matrix  $\mathbf{B}$  is diagonal, containing zeros everywhere except along the first  $r$  diagonal elements, where  $r$  is the rank of the matrix, with positive elements

$$\mathbf{B} = \begin{bmatrix} \sigma_1 & 0 & \cdot & \cdot \\ 0 & \sigma_2 & \cdot & \cdot \\ \cdot & \cdot & \ddots & \cdot \\ \cdot & \cdot & \cdot & \sigma_N \end{bmatrix} \quad (\text{F.2})$$

The columns of  $\mathbf{U}$  are the orthonormal eigenvectors of  $\mathbf{A}\mathbf{A}^+$ , while the columns of  $\mathbf{V}$  are the orthonormal eigenvectors of  $\mathbf{A}^+\mathbf{A}$ . The  $\sigma_j$  quantities are termed the singular values of  $\mathbf{A}$  and are the nonnegative square roots of the eigenvalues of  $\mathbf{A}^+\mathbf{A}$ . We will order them by decreasing magnitude:  $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_N \geq 0$ . Let  $\mathbf{A}$  be Hermitian with eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_N$ . Then the singular values of  $\mathbf{A}$  are  $|\lambda_1|, |\lambda_2|, \dots, |\lambda_N|$ . If  $\mathbf{A}$  is of rank  $r < N$ , then  $(N - r)$  of the  $\sigma_j$  terms are identically zero. The square of the “Frobenius norm” of  $\mathbf{A}$  is the sum of squares of the singular values of  $\mathbf{A}$ :

$$|\mathbf{A}|^2 = \sum_{j=1}^N \sigma_j^4 = \sum_{j=1}^N \lambda_j^2 \quad (\text{F.3})$$

With these formulas, the results in Section 4.1.1.5 follow at once.

**REFERENCES**

1. P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic Press, Orlando, FL, 1985.
2. H. C. Andrews and C. L. Patterson, *IEEE Trans. Comput.* **C-25**, 1976, 72.
3. S. C. Sahasrabudhe and A. D. Kulkarni, *Comput. Graph. Imag. Proc.* **9**, 1979, 203.

## APPENDIX G

# Derivation of Bounds for Degree of Polarization

I now present a simple derivation of the upper and lower bounds for the degree of polarization of a wave interacting with a linear optical system that is described by a Mueller matrix. My objective is to show that a general Mueller matrix deforms the surface  $\Sigma_2^{P_i}$  into an ellipsoid (see Section 4.1.2.5).

I assume that the field is described as a plane wave in terms of its Stokes vector

$$\mathbf{S} = \langle S_0 \rangle [1 \quad \langle \sigma_1 \rangle \quad \langle \sigma_2 \rangle \quad \langle \sigma_3 \rangle]^T = \langle S_0 \rangle [1 \quad \langle \sigma \rangle]^T \quad (\text{G.1})$$

where the three-dimensional vector  $\langle \sigma \rangle$  specifies the state of polarization of the wave. The degree of polarization can be written as

$$P = \left( \sum_{j=1}^3 \langle \sigma_j \rangle^2 \right)^{1/2} = |\langle \sigma \rangle| = |\mathbf{OM}| \quad (\text{G.2})$$

In the Stokes vector space the length of the vector  $\langle \sigma \rangle$  gives the degree of polarization of the wave, with the points  $O$  and  $M$  representing the origin of the unit ball  $\Sigma_1^3$  and the partially polarized state, respectively (see Section 3.2.2). We consider the surface  $\Sigma_2^P$  which defines the states of polarization where the degree of polarization is  $P$ . Note that  $\Sigma_2^1$  corresponding to  $|\langle \sigma \rangle| = 1$  is the Poincaré sphere, which is the locus of all possible pure states.

Let us write a general Mueller matrix as

$$\mathbf{M} = \begin{bmatrix} m_{00} & \mathbf{p}_1^T \\ \mathbf{p}_2 & \mathbf{n} \end{bmatrix} \quad (\text{G.3})$$

where  $\mathbf{p}_1$  and  $\mathbf{p}_2$  are  $3 \times 1$  real vectors and  $\mathbf{n}$  is a  $3 \times 3$  real matrix. The 0–0 entry  $m_{00}$  has been identified as being the gain of the optical system for incident unpolarized light. Substituting Eqs. (G.1) and (G.3) into Eq. (4.1.87) yields

$$1 = \frac{1}{g} (m_{00} + \mathbf{p}_1^T \cdot \langle \sigma \rangle_i) \quad (\text{G.4})$$

$$\langle \sigma \rangle_o = \frac{1}{g} (\mathbf{p}_2 + \mathbf{n} \langle \sigma \rangle_i) \quad (\text{G.5})$$

where  $g$  is the gain of the system. From Eqs. (G.4) and (G.5) we find

$$m_{00} = g_u \quad (G.6)$$

and

$$|\mathbf{p}_2| = m_{00}P_u \quad (G.7)$$

By the Minkowski inequality we have, from Eq. (G.5)

$$\frac{1}{g} |\mathbf{p}_2| - |\mathbf{n}\langle\sigma\rangle_i| \leq |\langle\sigma\rangle_o| = \frac{1}{g} |\mathbf{p}_2 + \mathbf{n}\langle\sigma\rangle_i| \leq \frac{1}{g} (|\mathbf{p}_2| + |\mathbf{n}\langle\sigma\rangle_i|) \quad (G.8)$$

If we substitute Eq. (G.7) in the latter expression, we get

$$\frac{1}{g} |m_{00}P_u - |\mathbf{n}\langle\sigma\rangle_i| \leq P_o \leq \frac{1}{g} (m_{00}P_u + |\mathbf{n}\langle\sigma\rangle_i|) \quad (G.9)$$

where  $|\mathbf{n}|$  denotes the matrix norm of  $\mathbf{n}$  and  $|\langle\sigma\rangle_i| = P_i$  is the degree of polarization of the incident wave. Observing that Eq. (G.4) yields bounds for  $g$  [i.e., Eqs. (4.1.97a, b)], we finally have

$$\frac{|m_{00}P_u - |\mathbf{n}\langle\sigma\rangle_i|}{m_{00} + |\mathbf{p}_1|P_i} \leq P_o \leq \frac{m_{00}P_u + |\mathbf{n}\langle\sigma\rangle_i|}{|m_{00} - |\mathbf{p}_1|P_i|} \quad (G.10)$$

Equation (G.9) is the central result of this appendix. We see that, in general,  $a \leq |\langle\sigma\rangle_o| \leq b$ , where the lower bound is nonnegative. Consequently, it turns out that  $\Sigma_2^{P_o}$  is, in the general case, an ellipsoid, which is the result we wished to prove. As a final remark, I wish to emphasize that the inequality (G.10) I have derived is valid for general Mueller matrices. A similar problem has been considered by Hovenier and van der Mee [1] and Lu and Chipman [2] for the case of Mueller-Jones matrices.

## REFERENCES

1. J. W. Hovenier and C. V. M. van der Mee, *Opt. Lett.* **20**, 1995, 2454.
2. Shi-Yau Lu and R. C. Chipman, *Opt. Commun.* (in press).

## Application of Maximum Entropy Principle in Polarization Optics

My purpose in this appendix is to discuss in some details an application of the maximum entropy principle (MEP) in the area of multiple scattering of light [1]. This result will be needed in Section 4.3.

It turns out that all information about an elastic scattering process is contained in the 16-element Mueller matrix. A number of restrictions are placed at the outset on the form of the  $\mathbf{M}$  matrix depending on the symmetry and reciprocity requirements. On one hand, the Mueller matrix should show Perrin symmetry that holds for elastic scattering from isotropic suspensions of particles, namely, in that case  $\mathbf{M}$  is diagonal for normal incidence. On the other hand, the optical medium is nondissipative. On introduction of these symmetries, the general form of the Mueller matrix  $\mathbf{M}$  can be written as

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & m_{11} & 0 & 0 \\ 0 & 0 & m_{22} & 0 \\ 0 & 0 & 0 & m_{33} \end{bmatrix} \quad (\text{H.1})$$

Now we evaluate analytically the elements  $m_{ii}$  ( $i = 1, 2, 3$ ) in Eq. (H.1) by an argument of maximum entropy. We proceed as follows. The entropy production per scattering reads as

$$\Delta S(n) = S(P(n + 1)) - S(P(n)) = h_s(n) \quad (\text{H.2})$$

where  $P(n)$  denotes the degree of polarization after  $n + 1$  scattering events. For the moment we shall not trouble with the precise form of  $h_s(n)$ . Here  $S(P)$  is given by Eq. (3.4.5), and the subscript  $s$  indicates that  $h_s$  depends on the particular state of polarization. With the help of Eq. (H.2), the total entropy production after  $n + 2$  scatterings reduces to

$$\Delta S = S(P(n + 1)) - S(P(0)) = \sum_{j=0}^n h_s(j) = \ln \left( \frac{s(P(0))}{s(P(n + 1))} \right) \quad (\text{H.3})$$

The function  $h_s(x)$  is taken to be a monotonically decreasing function from

$h_s(0)$  down to zero; this is to be expected from the theory of irreversible thermodynamics. We are seeking a candidate function  $h_s(x)$  in the metric space  $L^2$  that satisfies the condition  $h_s''(x) > 0$ , where the double prime indicates differentiation with respect to  $x$ . It is necessary to postulate the functional form of  $h_s(x)$ ; we have chosen to work out the function  $h_s(x) = a \exp(-bx)$ , which meets the requirements stated above.

Next, we consider an incident pure state of polarization that is linearly polarized [ $P_{n+1} = m_{11}(n)$ ]. From Eq. (H.3), we arrive at the relation

$$s(m_{11}) = \exp\left(-\sum_{j=0}^n h_s(j)\right) = \exp\left(-a\left(\frac{1 - \exp(-bn)}{1 - \exp(-b)}\right)\right) \quad (\text{H.4})$$

This is equivalent to saying that

$$m_{11}(n) = s^{-1}\left(A \sum_{j=0}^{n-1} B^j\right) \quad (\text{H.5})$$

where we have set for notational convenience  $A = \exp(-a)$  and  $B = \exp(-b)$ . We call attention to the important fact that  $B$  can be written from Eq. (H.2) as  $B = 1 - S(P(1))/\ln(2)$ ; consequently,  $B$  is fully determined by double scattering. Moreover, when maximum entropy is achieved (i.e., in the limit  $n \rightarrow \infty$ ), we require that  $A[1/(1 - B)] = \frac{1}{2}$ , that is,  $b = \ln(2)(1 - \exp(-a))$ . Putting everything together, we get the final expression for  $m_{11}$ :

$$m_{11}(n) = s^{-1}(2^{(B^n - 1)}) \quad (\text{H.6})$$

This equation allows the successive orders of iteration to be expressed in terms of the sole parameter  $B$ . It is worth noting that the same formula will apply for a pure state that is circularly polarized [i.e.,  $P_{n+1} = m_{33}(n)$ ] with a change in the value of  $B$ . Note that this method is quite general and may be used for a more involved Mueller matrix, but it does not tell us what kind of trial functions  $h_s(x)$  are to be used. This MEP simulation can incorporate any function  $h_s(x)$  that satisfies the physical constraints. With these formulas, the results in Section 4.4 follow at once.

## REFERENCE

1. D. Bicout and C. Brosseau, *J. Phys. I France* **2**, 1992, 2047.

## APPENDIX I

### Set of Coefficients $D_k^l$

$$D_0^0 = (|t_{11}|^2 + |t_{12}|^2 + |t_{21}|^2 + |t_{22}|^2) + (\langle |\tau_{11}|^2 \rangle + \langle |\tau_{12}|^2 \rangle + \langle |\tau_{21}|^2 \rangle + \langle |\tau_{22}|^2 \rangle) \quad (I.1)$$

$$D_1^0 = (|t_{11}|^2 - |t_{12}|^2 + |t_{21}|^2 - |t_{22}|^2) + (\langle |\tau_{11}|^2 \rangle - \langle |\tau_{12}|^2 \rangle + \langle |\tau_{21}|^2 \rangle - \langle |\tau_{22}|^2 \rangle) \quad (I.2)$$

$$D_2^0 = i(t_{11}^* t_{12} + t_{12}^* t_{11} + t_{21}^* t_{22} + t_{22}^* t_{21}) + i(\langle \tau_{11}^* \tau_{11} \rangle + \langle \tau_{22}^* \tau_{12} \rangle + \langle \tau_{21}^* \tau_{22} \rangle + \langle \tau_{22}^* \tau_{21} \rangle) \quad (I.3)$$

$$D_3^0 = i(t_{12}^* t_{11} - t_{11}^* t_{12} + t_{22}^* t_{21} - t_{21}^* t_{22}) + i(\langle \tau_{12}^* \tau_{11} \rangle - \langle \tau_{11}^* \tau_{12} \rangle + \langle \tau_{22}^* \tau_{21} \rangle - \langle \tau_{21}^* \tau_{22} \rangle) \quad (I.4)$$

$$D_0^1 = (|t_{11}|^2 + |t_{12}|^2 - |t_{21}|^2 - |t_{22}|^2) + (\langle |\tau_{11}|^2 \rangle + \langle |\tau_{12}|^2 \rangle - \langle |\tau_{21}|^2 \rangle - \langle |\tau_{22}|^2 \rangle) \quad (I.5)$$

$$D_1^1 = (|t_{11}|^2 - |t_{12}|^2 - |t_{21}|^2 + |t_{22}|^2) + (\langle |\tau_{11}|^2 \rangle - \langle |\tau_{12}|^2 \rangle + \langle |\tau_{21}|^2 \rangle - \langle |\tau_{22}|^2 \rangle) \quad (I.6)$$

$$D_2^1 = (t_{11}^* t_{12} + t_{12}^* t_{11} - t_{21}^* t_{22} - t_{22}^* t_{21}) + (\langle \tau_{11}^* \tau_{12} \rangle + \langle \tau_{12}^* \tau_{11} \rangle - \langle \tau_{22}^* \tau_{21} \rangle - \langle \tau_{21}^* \tau_{22} \rangle) \quad (I.7)$$

$$D_3^1 = i(t_{12}^* t_{11} - t_{11}^* t_{12} + t_{21}^* t_{22} - t_{22}^* t_{21}) + i(\langle \tau_{12}^* \tau_{11} \rangle - \langle \tau_{11}^* \tau_{12} \rangle + \langle \tau_{12}^* \tau_{22} \rangle - \langle \tau_{22}^* \tau_{12} \rangle) \quad (I.8)$$

$$D_0^2 = (t_{21}^* t_{11} + t_{11}^* t_{21} + t_{12}^* t_{22} + t_{22}^* t_{12}) + (\langle \tau_{21}^* \tau_{11} \rangle + \langle \tau_{11}^* \tau_{21} \rangle + \langle \tau_{12}^* \tau_{22} \rangle + \langle \tau_{22}^* \tau_{12} \rangle) \quad (I.9)$$

$$D_1^2 = (t_{21}^* t_{11} + t_{11}^* t_{21} - t_{12}^* t_{22} - t_{22}^* t_{12}) + (\langle \tau_{21}^* \tau_{11} \rangle + \langle \tau_{11}^* \tau_{21} \rangle - \langle \tau_{12}^* \tau_{22} \rangle - \langle \tau_{22}^* \tau_{12} \rangle) \quad (I.10)$$

$$D_2^2 = (t_{11}^* t_{22} + t_{22}^* t_{11} + t_{21}^* t_{12} + t_{12}^* t_{21}) + (\langle \tau_{11}^* \tau_{22} \rangle + \langle \tau_{22}^* \tau_{11} \rangle + \langle \tau_{21}^* \tau_{12} \rangle + \langle \tau_{12}^* \tau_{21} \rangle) \quad (I.11)$$

$$D_3^2 = i(t_{22}^* t_{11} - t_{11}^* t_{22} + t_{21}^* t_{12}^* + t_{12}^* t_{21}) + i(\langle \tau_{22}^* \tau_{11} \rangle - \langle \tau_{11}^* \tau_{22} \rangle \\ + \langle \tau_{12}^* \tau_{21} \rangle - \langle \tau_{21}^* \tau_{12} \rangle) \quad (I.12)$$

$$D_0^3 = i(t_{11}^* t_{21} - t_{21}^* t_{11} + t_{12}^* t_{22} - t_{22}^* t_{12}) + i(\langle \tau_{11}^* \tau_{21} \rangle - \langle \tau_{21}^* \tau_{11} \rangle \\ + \langle \tau_{12}^* \tau_{22} \rangle - \langle \tau_{22}^* \tau_{12} \rangle) \quad (I.13)$$

$$D_1^3 = i(t_{11}^* t_{21} - t_{21}^* t_{11} + t_{22}^* t_{12} - t_{12}^* t_{22}) + i(\langle \tau_{11}^* \tau_{21} \rangle - \langle \tau_{21}^* \tau_{11} \rangle \\ + \langle \tau_{22}^* \tau_{12} \rangle - \langle \tau_{12}^* \tau_{22} \rangle) \quad (I.14)$$

$$D_2^3 = i(t_{11}^* t_{22} - t_{22}^* t_{11} + t_{12}^* t_{21} - t_{21}^* t_{12}) + i(\langle \tau_{11}^* \tau_{22} \rangle - \langle \tau_{22}^* \tau_{11} \rangle \\ + \langle \tau_{12}^* \tau_{21} \rangle - \langle \tau_{21}^* \tau_{12} \rangle) \quad (I.15)$$

$$D_3^3 = (t_{11}^* t_{22} + t_{22}^* t_{11} - t_{21}^* t_{12} - t_{12}^* t_{21}) + (\langle \tau_{11}^* \tau_{22} \rangle + \langle \tau_{22}^* \tau_{11} \rangle \\ - \langle \tau_{21}^* \tau_{12} \rangle - \langle \tau_{12}^* \tau_{21} \rangle) \quad (I.16)$$

## APPENDIX J

# Evaluation of Correlation Function in Eq. (4.4.34)

In this appendix, we evaluate the correlation function of the field in Eq. (4.4.34) of the text. To proceed, we use the Bethe–Salpeter (BS) equation formalism. We limit our discussion to transmission of a plane wave that is normally incident on a medium with randomly positioned isotropic pointlike scatterers. In the analysis about to follow, we first review the basic equations that will be useful in our discussion. This involves introducing a number of important notions that are hardly in the stock in trade of optical physicists. We shall only give the briefest account here. For further information, the reader should consult Refs. 1–4.

Our starting point is the set of Maxwell's equations for the electric field strength at the point  $\mathbf{r}$ :

$$\nabla^2 E_i + k_0^2(1 + u(\mathbf{R}))\delta_{ij}E_j = J_i \quad (\text{J.1a})$$

$$\nabla \cdot \mathbf{E} = 0 \quad (\text{J.1b})$$

where  $J_i$  denotes the density of the radiation sources and  $u(\mathbf{R})$  represents the fluctuating part of the permittivity as a function of the position vector  $\mathbf{R}$ . We have omitted a term  $\nabla(\mathbf{E} \cdot \nabla \ln(1 + u(\mathbf{R})))$ , which is small if the variation of disorder is slow compared with the wavelength of the light. The basic field variable is represented by its complex analytic signal. We can solve for the field  $\mathbf{E}$  in terms of the source  $\mathbf{J}$  in terms of Green's function. Following this approach, Eq. (J.1a) can be rewritten into an integral equation that expresses the linear response of the scattering medium to incident light:

$$E_i(\mathbf{r}) = \int d^3\mathbf{R}' G_{ij}(\mathbf{R}, \mathbf{R}') J_j(\mathbf{R}') \quad (\text{J.2})$$

where Green's function  $G(\mathbf{R}, \mathbf{R}')$  describes the effect of a point source localized at position  $\mathbf{R}'$  on the observation point  $\mathbf{R}$ . In (J.2),  $G_{ij}$  denotes the single Green function of (J.1a); it obeys the reciprocity relation,  $G_{ij}(\mathbf{R}, \mathbf{R}') = G_{ji}(\mathbf{R}', \mathbf{R})$ , which states that, in an arbitrary linear inhomogeneous medium, the field remains unchanged if the source location and observation point are

interchanged. The Green function measures the local response of the electric field at a given point to a local disturbance of the electric field at another given point.

Now, we turn to the average Green function. In free space,  $u(\mathbf{R}) \equiv 0$ , the averaged Green function for a scalar wave in direct space is usually written as

$$\langle G(\mathbf{R}, \mathbf{R}') \rangle = \langle G(\mathbf{R} - \mathbf{R}') \rangle = -\frac{\exp\left(\left(ik_o - \frac{1}{2l}\right)|\mathbf{R} - \mathbf{R}'|\right)}{4\pi|\mathbf{R} - \mathbf{R}'|} \quad (J.3)$$

where the scattering mean free path  $l$  describes the decay length of the Green function and  $k_o$  is the vacuum wavenumber. Note that as a result of macroscopic homogeneity, the function  $G(\mathbf{R}, \mathbf{R}')$  averaged over disorder has translational invariance, that is, it depends on  $|\mathbf{R} - \mathbf{R}'|$  only. Its Fourier transform is expressed as

$$\langle G(q) \rangle = \left( k_o^2 - q^2 - i \frac{k_o}{l} \right)^{-1} \quad (J.4)$$

For polarized waves, the averaged Green function of (J.1) is now of the form

$$\langle G_{ik}(\mathbf{q}) \rangle = (\delta_{ik} - \mathbf{q}_i \mathbf{q}_k) \langle G(\mathbf{q}) \rangle \quad (J.5)$$

where  $\mathbf{q}$  denotes a unit vector. The following calculations are more conveniently done in momentum space. Up to this point our various formal manipulations have been exact. Although some progress is possible, it proves convenient at this stage to introduce an approximation. The theory that follows from this assumption is known as the “ladder” approximation. The BS equation derived for the second-order Green function can be obtained from the usual diagrammatic expansion

$$\langle G_{im} G_{jn}^* \rangle = G_{im} G_{jn}^* + G_{im'} G_{jn'}^* W_{ijm'n'} G_{m'm} G_{n'n}^* \quad (J.6)$$

where summation over repeated indices is presumed, with  $G^*$  the complex conjugate of  $G$ . The first term on the right side of (J.6) represents the coherent contribution, which goes to zero very rapidly and is neglected, while the other term is the incoherent contribution. The BS equation under the ladder approximation of uncorrelated discrete scatterers results in the usual vector radiative transfer equation [5–7].

In the weak scattering situation, the observed intensity is an incoherent sum of contributions of light scattered through all possible paths. For a polarized field, the evaluation of the intensity operator  $W_{ijm'n'}$  is more involved than for the scalar situation. For scalar waves, one shows that within the diffusion approximation (i.e., incoherent and diffusive transport of light), the incoherent

intensity can be written as

$$I_{\text{inc}}(\mathbf{R}) = \frac{l^2}{36\pi^2} \sum_q W(\mathbf{q}) \exp(i\mathbf{q} \cdot \mathbf{R}) = \frac{6\pi}{l} \sum_{n=1}^{\infty} F(\mathbf{R}, n) \quad (\text{J.7})$$

where the Green function  $F(\mathbf{R}, n)$  of the diffusion equation is the weight of diffusion paths and may be obtained by summing successive ladder diagrams.

Now, the intensity operator for a polarized wavefield can be written as

$$W_{ijmn}(\mathbf{q}) = \frac{6\pi}{l\Omega} \sum_{n=0} \sum_a \langle ij | a \rangle (\lambda_a(\mathbf{q}))^n \langle a | mn \rangle \quad (\text{J.8})$$

The first term  $n = 0$  corresponds to the single-scattering situation

$$W_{ijmn}^{(0)}(\mathbf{q}) = \frac{6\pi}{l\Omega} \delta_{im} \delta_{jn} \quad (\text{J.9})$$

The nine orthonormal eigenvectors  $|a\rangle$  and associated eigenvalues  $\lambda_a(\mathbf{q})$  are given, to this degree of approximation [i.e.,  $O(q^2)$ ], by [1, 2]

$$|1\rangle = \frac{1}{\sqrt{3}} (|11\rangle + |22\rangle + |33\rangle); \quad \lambda_1 = 1 - \frac{q^2 l^2}{3} \quad (\text{J.10a})$$

$$|2\rangle = \frac{1}{\sqrt{3}} (|11\rangle + b|22\rangle + b^2|33\rangle); \quad \lambda_2 = \frac{7}{10} \left(1 - \frac{q^2 l^2}{3}\right) \quad (\text{J.10b})$$

$$|3\rangle = \frac{1}{\sqrt{3}} (|11\rangle + b^*|22\rangle + b^{*2}|33\rangle); \quad \lambda_3 = \frac{7}{10} \left(1 - \frac{q^2 l^2}{3}\right) \quad (\text{J.10c})$$

$$|4\rangle = \frac{1}{\sqrt{2}} (|12\rangle + |21\rangle); \quad \lambda_4 = \frac{7}{10} - \frac{23}{70} q^2 l^2 + \frac{l^2}{7} (\mathbf{q}_1^2 + \mathbf{q}_2^2) \quad (\text{J.10d})$$

$$|5\rangle = \frac{1}{\sqrt{2}} (|13\rangle + |31\rangle); \quad \lambda_5 = \frac{7}{10} - \frac{23}{70} q^2 l^2 + \frac{l^2}{7} (\mathbf{q}_1^2 + \mathbf{q}_3^2) \quad (\text{J.10e})$$

$$|6\rangle = \frac{1}{\sqrt{2}} (|23\rangle + |32\rangle); \quad \lambda_6 = \frac{7}{10} - \frac{23}{70} q^2 l^2 + \frac{l^2}{7} (\mathbf{q}_2^2 + \mathbf{q}_3^2) \quad (\text{J.10f})$$

$$|7\rangle = \frac{1}{\sqrt{2}} (|12\rangle - |21\rangle); \quad \lambda_7 = \frac{1}{2} - \frac{3}{10} q^2 l^2 + \frac{l^2}{5} (\mathbf{q}_1^2 + \mathbf{q}_2^2) \quad (\text{J.10g})$$

$$|8\rangle = \frac{1}{\sqrt{2}} (|13\rangle - |31\rangle); \quad \lambda_8 = \frac{1}{2} - \frac{3}{10} q^2 l^2 + \frac{l^2}{5} (\mathbf{q}_1^2 + \mathbf{q}_3^2) \quad (\text{J.10h})$$

$$|9\rangle = \frac{1}{\sqrt{2}} (|23\rangle - |32\rangle); \quad \lambda_9 = \frac{1}{2} - \frac{3}{10} q^2 l^2 + \frac{l^2}{5} (\mathbf{q}_2^2 + \mathbf{q}_3^2) \quad (\text{J.10i})$$

where we have set for notational convenience  $b = e^{2i\pi/3}$ . At the present level of approximation, one finds from (J.7) and (J.8)

$$\langle G_{im} G_{jn}^* \rangle = \frac{3}{2} \sum_{a=1}^9 \langle ij | a \rangle \lambda_a(\mathbf{R}, n) \langle a | mn \rangle \quad (\text{J.11})$$

with

$$\lambda_a(\mathbf{R}, n) = \frac{1}{\Omega} \sum_q (\lambda_a(q))^n \exp(iq \cdot \mathbf{R}) \quad (\text{J.12})$$

where  $|ij\rangle$ ,  $|mn\rangle$  denote respectively the initial and final states of polarization. Equations (J.10) contain at least one thing worth mentioning. We note that  $1 - \lambda_1(0)$  vanishes with  $q \rightarrow 0$ . It turns out that the mode with eigenvalue  $\lambda_1(\mathbf{q})$  corresponds to the Goldstone mode for this problem (i.e., appearance of infinite range correlation). This mode originates from the long-range diffusion, while the eigenvalues  $\lambda_2, \dots, \lambda_9$  are associated with short paths.

Since we are interested in evaluating the dependence of the mean intensity as a function of the scattering number, we write

$$\langle G_{im} G_{jn}^* \rangle = F_{ijmn}(\mathbf{R}, n) \quad (\text{J.13})$$

As the direction of propagation is along  $\mathbf{e}_3$ , we have

$$\langle G_{11} G_{11}^* \rangle = \langle G_{22} G_{22}^* \rangle = F_{2222} = \frac{1}{2}[1 + 2(\frac{7}{10})^n]F(\mathbf{R}, n, L_1) \quad (\text{J.14a})$$

$$\langle G_{12} G_{12}^* \rangle = F_{1122} = \frac{1}{2}[1 - (\frac{7}{10})^n]F(\mathbf{R}, n, L_1) \quad (\text{J.14b})$$

$$\langle G_{11} G_{22}^* \rangle = F_{1212} = \frac{3}{4}[(\frac{7}{10})^n F(\mathbf{R}, n, L_2) + (\frac{1}{2})^n F(\mathbf{R}, n, L_3)] \quad (\text{J.14c})$$

$$\langle G_{12} G_{21}^* \rangle = F_{1221} = \frac{3}{4}[(\frac{7}{10})^n F(\mathbf{R}, n, L_2) - (\frac{1}{2})^n F(\mathbf{R}, n, L_3)] \quad (\text{J.14d})$$

with the Green function  $F(\mathbf{R}, n, L_i)$  for an infinite medium given by

$$F(\mathbf{R}, n, L_i) = \left( \frac{1}{4\pi L_i^2 n} \right)^{3/2} \exp \left\{ - \frac{R^2}{4L_i^2 n} \right\} \quad (\text{J.15})$$

where we have set  $L_1^2 = l^2/3$ ,  $L_2^2 = 23l^2/49$ , and  $L_3^2 = 3l^2/5$ . It follows at once that carrying out the integration of Eqs. (J.14) over the entire space, the  $\langle G_{im} G_{jn}^* \rangle$  are given by the formulas:

$$\langle G_{11} G_{11}^* \rangle = \langle G_{22} G_{22}^* \rangle = \frac{1}{2}[1 + 2(\frac{7}{10})^n] \quad (\text{J.16a})$$

$$\langle G_{12} G_{12}^* \rangle = \frac{1}{2}[1 - (\frac{7}{10})^n] \quad (\text{J.16b})$$

$$\langle G_{11} G_{22}^* \rangle = \frac{3}{4}[(\frac{7}{10})^n + (\frac{1}{2})^n] \quad (\text{J.16c})$$

$$\langle G_{12} G_{21}^* \rangle = \frac{3}{4}[(\frac{7}{10})^n - (\frac{1}{2})^n] \quad (\text{J.16d})$$

We see that Eq. (4.4.35) follows from Eq. (4.4.34) when Eqs. (J.16a-d) are used.

**REFERENCES**

1. M. J. Stephen and G. Cwilich, *Phys. Rev.* **B34**, 1986, 7564. See also M. J. Stephen, *Phys. Rev.* **B37**, 1988, 1.
2. F. C. Macintosh and S. John, *Phys. Rev.* **B40**, 1989, 2383.
3. Y. N. Barabanenkov, *Sov. Phys. Usp.* **18**, 1975, 673.
4. J. K. Lee and J. A. Kong, *J. Opt. Soc. Am. A2*, 1985, 2171.
5. E. Akkermans, P. E. Wolf, R. Maynard, and G. Maret, *J. Phys. France* **49**, 1988, 77.
6. D. Bicout, E. Akkermans, and R. Maynard, *J. Phys. I France* **1**, 1991, 471.
7. D. Bicout and C. Brosseau, *J. Phys. I France* **2**, 1992, 2047.

## APPENDIX K

# Diffusing-Wave Spectroscopy

Measurement of the autocorrelation function of the temporal fluctuations in the intensity of multiply scattered light by the motion of the scattering centres is termed *diffusing-wave spectroscopy* (DWS). This technique is described in detail in the literature and was devised to extend the usual quasielastic light scattering to the multiple-scattering regime for studying the positions and motions of scatterers in optically dense systems that multiply scatter the probing light [1, 2]. For particles that are undergoing random motion, such as, Brownian motion, the phases of the individual scattered waves are changing randomly and independently of the other scattered waves. The intensity at the detector will thus fluctuate. The timescale of the intensity fluctuations is related to the rate at which the phase of the scattered waves is changing and thus depends on the motion of the scattering particles. By monitoring the intensity fluctuations, it is possible to derive information about the motion of the scattering particles. Usually one measures the normalized temporal intensity autocorrelation function  $g^{(2)}(t) = \langle S_0(t)S_0(0) \rangle / \langle S_0(0) \rangle^2$ , where  $S_0(t)$  is the light intensity at time  $t$ . The angle brackets denote an ensemble average. For an ergodic system, this ensemble average is equivalent to a time averaged over a time long compared with the coherence time  $\tau_2$  of the light source. Most calculations of the effects of sample fluctuations on scattered light yield expressions for the normalized temporal electric field autocorrelation function,  $g^{(1)}(t) = \langle E(t)E^*(0) \rangle / \langle S_0(0) \rangle$ . For a zero-mean complex Gaussian process,  $g^{(1)}(t)$  and  $g^{(2)}(t)$  are related through the Siegert-Reed factorization theorem:  $g^{(2)}(t) \approx 1 + |g^{(1)}(t)|^2$  [3]. Maret and Wolf related the measurement of the field autocorrelation function  $g^{(1)}(t)$  to the Green function describing the diffusive transport of the intensity [1]. The form of these auto correlation functions depends on the state of polarization. As was evidenced by Mackintosh et al., each diffusion path leads to a decay of  $g^{(1)}(t)$ , which depends on the scattering events [4]. For very short times, the correlation function is dominated by long paths for which the waves have lost memory of their incident state of polarization.

In the following, we will derive closed-form expressions of the degree of polarization for multiply scattered light from a half-space. For the incident pure state of polarization  $[\langle S_0 \rangle \quad \langle S_1 \rangle \quad \langle S_2 \rangle \quad \langle S_3 \rangle]^T$ , it can be shown that

the output degree of polarization is expressed (for  $t^* < 1$ ) as

$$P_o(t^*) = \{\langle S_1 \rangle^2 P_1^2(t^*) + \langle S_2 \rangle^2 P_2^2(t^*) + \langle S_3 \rangle^2 P_3^2(t^*)\}^{1/2} \quad (\text{K.1})$$

where the  $P_j(t)$  values denote respectively the degree of polarization for linear polarization, linear polarization at  $45^\circ$ , and circular polarization;  $t^* \equiv t/\tau_o$  is the dimensionless time; and  $\tau_o$  is the time required by a scatterer to move one optical wavelength [5].

The motivation for this appendix is the calculation of the degree of polarization for multiply scattered light in the context of DWS. For scattering particles executing simple Brownian random walks with a diffusion coefficient  $D_B$ , we can show that in the weak scattering limit and within the diffusion approximation, the scattered temporal field correlation function is given by

$$g^{(1)}(\mathbf{R}, t) = \frac{\langle E(0)E^*(t) \rangle}{\langle |E(0)|^2 \rangle} = \frac{\int F(\mathbf{R}, n) \exp(-2nt/\tau_o) dn}{\int F(\mathbf{R}, n) dn} \quad (\text{K.2})$$

where  $F(\mathbf{R}, n)$ , denoting the Green function of the diffusion equation, is the number of scattering paths of length  $nl$ ,  $\tau_o = k_o^{-2}D_B^{-1}$  is the diffusion time of a scatterer over a distance given by the wavelength of the light, and  $k_o$  is the lightwavenumber in the medium. From inspection of Eq. (K.2), we note that  $g^{(1)}(\mathbf{R}, t)$  is the normalized Laplace transform of  $F(\mathbf{R}, n)$ , thus giving a one-to-one correspondence between the number of scattering events  $n$  and the dimensionless variable  $t^* = t/\tau_o$ .

Now we consider the special situation of reflected light from a homogeneous nonabsorbing half-space. Let us choose a right-hand Cartesian coordinate system of axes  $(x_1, x_2, x_3)$ . In this geometry, we assume that the scattering medium occupies the region of space  $x_3 > 0$  and that the light is incident normally on the semiinfinite medium from the vacuum  $x_3 = -\infty$ . For such semiinfinite medium, the weight of diffusion paths for scalar waves may be obtained from the preceding infinite-space case by the method of images. For this system, the integral in (K.2) can be easily evaluated when a scalar field is considered from results of Appendix J. As a result, Eq. (K.2) reduces to a stretched exponential [1, 6]

$$g^{(1)}(t) = \frac{\int_1^\infty F(n) \exp(-2nt^*) dn}{\int_1^\infty F(n) dn} = \exp\{-\sqrt{6t^*}\} \quad (\text{K.3})$$

For a polarized field, we need to specify the particular state of polarization:

1. *Linear Polarization.* Considering an incident field polarized along  $\mathbf{e}_1$  (labeled  $\parallel$ ), we get

$$g_{\parallel}^{(1)}(t) = \frac{\langle E_1(0)E_1^*(t) \rangle}{\langle |E_1(0)|^2 \rangle} = \frac{\int F_{\parallel}(n) \exp(-2nt^*)dn}{\int F_{\parallel}(n)dn} \quad (\text{K.4a})$$

Similarly for the  $\mathbf{e}_2$  component (labeled  $\perp$ ), we have

$$g_{\perp}^{(1)}(t) = \frac{\langle E_2(0)E_2^*(t) \rangle}{\langle |E_2(0)|^2 \rangle} = \frac{\int F_{\perp}(n) \exp(-2nt^*)dn}{\int F_{\perp}(n)dn} \quad (\text{K.4b})$$

where it follows from results of Appendix J that

$$F_{\parallel}(n) = \langle 11|F|11 \rangle = \langle G_{11}G_{11}^* \rangle = \frac{1}{2}[1 + 2(\frac{7}{10})^n]F(n, L_1) \quad (\text{K.5a})$$

$$F_{\perp}(n) = \langle 11|F|22 \rangle = \langle G_{12}G_{12}^* \rangle = \frac{1}{2}[1 - (\frac{7}{10})^n]F(n, L_1) \quad (\text{K.5b})$$

The result is

$$g_{\parallel}^{(1)}(t) = \frac{\exp(-a) + 2\exp(-b)}{1 + 2\exp(-b')} \quad (\text{K.6a})$$

$$g_{\perp}^{(1)}(t) = \frac{\exp(-a) - \exp(-b)}{1 - \exp(-b')} \quad (\text{K.6b})$$

where  $a = \sqrt{6t^*}$ ;  $b = \sqrt{6t^* + 3\ln(\frac{10}{7})}$ , and  $b' = \sqrt{3\ln(\frac{10}{7})}$ . The degree of polarization is then given by a simple relation

$$P_1(t) = \frac{|g_{\parallel}^{(1)}(t) - g_{\perp}^{(1)}(t)|}{g_{\parallel}^{(1)}(t) + g_{\perp}^{(1)}(t)} = \frac{N}{D} \quad (\text{K.7})$$

2. *Circular Polarization.* In a similar manner, the case of circular polarization is developed. The incident unit vector is  $\mathbf{e}_r = 1/\sqrt{2}(\mathbf{e}_1 - i\mathbf{e}_2)$ , and the polarization states may be written as

$$|ll\rangle = \frac{1}{2}\{|11\rangle - i|21\rangle + i|12\rangle + |22\rangle\} \quad (\text{K.8a})$$

$$|lr\rangle = \frac{1}{2}\{|11\rangle + i|21\rangle - i|12\rangle + |22\rangle\} \quad (\text{K.8b})$$

In passing, we would like to remark on the permutation symmetry of these vectors under the conjugation operation. We call  $F_+$  (resp.  $F_-$ ) the Green function corresponding to the positive (resp. negative) helicity. In like fashion as above, we have

$$F_+ = \langle ll|F|ll\rangle \quad (\text{K.9a})$$

$$F_- = \langle lr|F|ll\rangle \quad (\text{K.9b})$$

with

$$\begin{aligned} F_+ \\ F_- \end{aligned} \left\{ \right. = \frac{1}{2} \{ \langle G_{11}G_{11}^* \rangle + \langle G_{12}G_{12}^* \rangle \pm \langle G_{11}G_{22}^* \rangle - (\pm) \langle G_{12}G_{21}^* \rangle \} \quad (\text{K.10a})$$

and

$$\begin{aligned} F_+(n) \\ F_-(n) \end{aligned} \left\{ \right. = \frac{1}{4} [2 + (\frac{7}{10})^n \pm \frac{5}{3}(\frac{1}{2})^n] F(n, L_1) \quad (\text{K.10b})$$

The corresponding temporal correlation functions are

$$\begin{aligned} g_r^{(1)}(t) \\ g_l^{(1)}(t) \end{aligned} \left\{ \right. = \frac{2 \exp(-a) + \exp(-b) \pm \frac{5}{3} \exp(-d)}{2 + \exp(-b') \pm \frac{5}{3} \exp(-d')} \quad (\text{K.11})$$

with  $a, b, b'$  given by the aforementioned expressions and  $d = \sqrt{(\frac{5}{3})(2t^* + \ln(2))}$ ;  $d' = \sqrt{(\frac{5}{3})\ln(2)}$ . The corresponding degree of polarization is expressed as

$$P_3(t) = \frac{|g_l^{(1)}(t) - g_r^{(1)}(t)|}{g_l^{(1)}(t) + g_r^{(1)}(t)} = \frac{N}{D} \quad (\text{K.12})$$

with  $\frac{D}{N} \left\{ \right. = B_1 B_3 \pm B_2 B_4$ , and

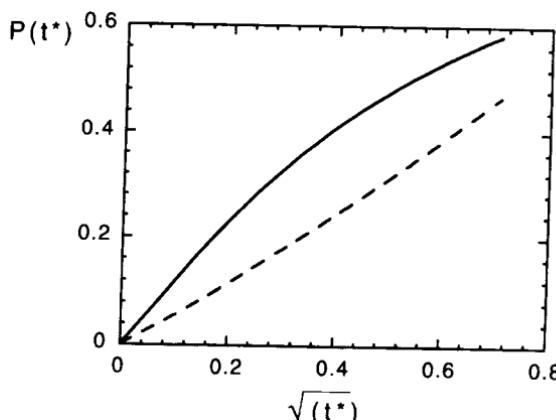
$$B_1 = 2 \exp(-a) + \exp(-b) + \frac{5}{3} \exp(-d) \quad (\text{K.13a})$$

$$B_2 = 2 \exp(-a) + \exp(-b) - \frac{5}{3} \exp(-d) \quad (\text{K.13b})$$

$$B_3 = 2 + \exp(-b') + \frac{5}{3} \exp(-d') \quad (\text{K.13c})$$

$$B_4 = 2 + \exp(-b') - \frac{5}{3} \exp(-d') \quad (\text{K.13d})$$

As displayed in Fig. K.1, the light is completely depolarized at very short times ( $t^* \ll 1$  corresponds to long diffusion paths). Because of the different contributions due to long paths, the  $P_o(t^*)$  versus  $(t^*)^{1/2}$  differ in slope according to the input state of polarization. As before, we find that complete depolarization of circularly polarized incident light requires less scattering events than for the



**Fig. K.1.** Polarization dependence of the autocorrelation functions Eq. (J.1) as a function of the dimensionless time  $t^*$ . The curves represent linear parallel polarization (solid line) and right circular polarization (dashed line).

linearly polarized case. This contrasts with longer times (short diffusion paths) [5].

These experiments hold enormous potential allowing detailed investigation of the hydrodynamics of dense complex fluids such as gels and colloids, in the regime of high concentration, (i.e.,  $d \gg l^*$ ) so that traditional scattering theory based on the single-scattering (Born) approximation is no longer valid. In closing, it is also worth mentioning that DWS may be useful to probe noninvasively heterogeneous flow and turbulence in turbid media. Other potential applications of DWS are in connection with the development of new methods for imaging of tumors and tomography as well as for biological tissue characterization [7].

## REFERENCES

1. G. Maret and P. E. Wolf, *Z. Phys.* **B65**, 1987, 409.
2. M. Rosenbluh, M. Hoshen, I. Freund, and M. Kaveh, *Phys. Rev. Lett.* **58**, 1987, 2754.
3. A. J. F. Siegert, MIT Rad. Lab. Report No. 465, 1943. See also I. S. Reed, *IRE Trans. Inf. Theory* **IT-8**, 1962, 194.
4. F. C. Mackintosh and S. John, *Phys. Rev.* **B40**, 1989, 2383.
5. D. Bicout and C. Brosseau, *J. Phys. I France* **2**, 1992, 2047.
6. D. J. Pine, D. A. Weitz, P. M. Chaikin, and E. Herbolzheimer, *Phys. Rev. Lett.* **60**, 1988, 1134.
7. See the feature issue titled "Diffusing photons in turbid media," *J. Opt. Soc. Am. A14*, 1997.

## APPENDIX L

---

# Polarization of Photons

One of the questions that baffled the great thinkers in the early 1900s was: "When is a quantized field description of light necessary?" There has been, since the earliest days of quantum mechanics, a large body of work toward answering this question. We shall not attempt to discuss these works in much generality since our interest here is to extend the classical description of partial polarization to quantum electrodynamics. Up to now we have discussed the description of polarization of a radiation field purely from the point of view of classic electrodynamics. So far we have worked exclusively with waves and have avoided talking about photons. The particle-like manifestations of the electromagnetic field are described quantum-mechanically in terms of photons. We may expect the dynamics of the electromagnetic field to be governed by quantum mechanics, as is the dynamics of particles. A much greater variety of states is available to quantized fields than to classic ones, including the antibunching and the squeezing of field noise [1]. For a more complete discussion of the quantum theory of light, the reader may consult, for example, Refs. [1–3].

In Section 3.2.5 we gave a detailed discussion of the properties of partially polarized beams of plane electromagnetic waves based on an analogy with spin- $\frac{1}{2}$  systems. However, a photon is not a spin- $\frac{1}{2}$  system: it is a massless quantized particle with spin-1 but of a special kind in that, because of the transversality of electromagnetic waves, the spin component along the direction of propagation behaves as a two-valued (instead of a three-valued) degree of freedom.

Let us begin our development by reviewing a few facts about ordinary quantum field theory. In Section 2.3.4 we learned that the classic description of the electromagnetic field in free space, in terms of plane waves, rests on an infinite set of independent modes, each analogous to harmonic oscillators. As in classic physics, the simple harmonic oscillator occupies a privileged position in quantum physics. From the quantal point of view, the field variables at one particular point in space are the photon creation and annihilation operators. Then the quantized electromagnetic field behaves as a system of bosons. To fix notation, we now define the appropriate physical quantities with which we will deal. By way of background, we first briefly describe the two-mode description

of the field operator, which can be understood as the analog of the decomposition of the classic field in two orthogonal components [i.e., Eq. (3.1.1)]. In parametrizing quantized fields, it is convenient to separate the electromagnetic field as a sum of right- and left-traveling waves as

$$\mathbf{E}(\mathbf{R}, t) = \mathbf{E}^+(\mathbf{R}, t) + \mathbf{E}^-(\mathbf{R}, t) \quad (\text{L.1})$$

The positive-frequency part of the electromagnetic field  $\mathbf{E}^+$  contains all terms that vary as  $\exp(-2i\pi\nu t)$ , whereas the negative-frequency part  $\mathbf{E}^-$  contains all those that vary as  $\exp(2i\pi\nu t)$  and is the Hermitian adjoint of  $\mathbf{E}^+$ . These two terms describe the absorption and emission, respectively, of a single photon at point  $\mathbf{R}$  and time  $t$ . To make things well defined, we write the transverse field operator in the linear polarization basis as

$$\mathbf{E}(\mathbf{R}, t) = i \sum_k \left( \frac{1}{2} \frac{\hbar v_k}{\epsilon_0 V} \right)^{1/2} (\mathbf{a}_k \mathbf{u}_k(\mathbf{R}) \exp(-2i\pi\nu t) - \mathbf{a}_k^+ \mathbf{u}_k^*(\mathbf{R}) \exp(2i\pi\nu t)) \quad (\text{L.2})$$

where  $k$  denotes the normal modes of the field in a box of volume  $V$ . As we have seen in Section 2.3, their number is denumerable if we think of the field as enclosed in a finite volume. The functions  $\mathbf{u}_k(\mathbf{R})$  form an orthonormal set of vector mode functions, and the  $\nu_k$  are the corresponding frequencies;  $\mathbf{a}_k$  is the annihilation operator for the  $k$ th mode, and its adjoint,  $\mathbf{a}_k^+$ , is the corresponding creation operator. These operators satisfy the familiar bosonic commutation rules

$$[\mathbf{a}_k, \mathbf{a}_{k'}^+] = \delta_{kk'} \quad (\text{L.3})$$

$$[\mathbf{a}_k, \mathbf{a}_{k'}] = [\mathbf{a}_k^+, \mathbf{a}_{k'}^+] = 0 \quad (\text{L.4})$$

Up to this point we have considered pure quantum states. A field in the form of a statistical mixture is described by a density operator  $\mathbf{D}$  that is Hermitian, positive definite, and of trace one. With the help of this operator, the statistical average of any observable  $\mathbf{O}$  is obtained as  $\langle \mathbf{O} \rangle = \text{tr}(\mathbf{OD})$ , in a similar fashion as was done for a classic theory of coherence.

Attention is now turned to the analysis of field correlations. The Stokes parameters, which are real numbers in classic electrodynamics, become operators in quantum electrodynamics. From the two-mode description of the field, we define the following Hermitian Stokes operators  $\Sigma_j$ :

$$\Sigma_0 = \mathbf{a}_1^+ \mathbf{a}_1 + \mathbf{a}_2^+ \mathbf{a}_2 \quad (\text{L.5a})$$

$$\Sigma_1 = \mathbf{a}_1^+ \mathbf{a}_1 - \mathbf{a}_2^+ \mathbf{a}_2 \quad (\text{L.5b})$$

$$\Sigma_2 = \mathbf{a}_2^+ \mathbf{a}_1 + \mathbf{a}_1^+ \mathbf{a}_2 \quad (\text{L.5c})$$

$$\Sigma_3 = i(\mathbf{a}_2^+ \mathbf{a}_1 - \mathbf{a}_1^+ \mathbf{a}_2) \quad (\text{L.5d})$$

The helicity basis can be used alternatively to describe the polarization properties of photons. The result is

$$\Sigma_0^{(lr)} = \mathbf{a}_r^+ \mathbf{a}_r + \mathbf{a}_l^+ \mathbf{a}_l \quad (\text{L.6a})$$

$$\Sigma_1^{(lr)} = \mathbf{a}_l^+ \mathbf{a}_r + \mathbf{a}_r^+ \mathbf{a}_l \quad (\text{L.6b})$$

$$\Sigma_2^{(lr)} = i(\mathbf{a}_r^+ \mathbf{a}_l - \mathbf{a}_l^+ \mathbf{a}_r) \quad (\text{L.6c})$$

$$\Sigma_3^{(lr)} = \mathbf{a}_l^+ \mathbf{a}_l - \mathbf{a}_r^+ \mathbf{a}_r \quad (\text{L.6d})$$

The Stokes operators contain all the relevant information about the polarization of the field. The connection between circular polarization and photon longitudinal spin polarization is established by noting that the photon helicity is given by the operator  $-i\Sigma_3$  [1]. Apart from a factor of 2, one can check that the Stokes operators  $\Sigma_j$  ( $j = 1, 2, 3$ ) satisfy the same commutation rules as does the angular moment while  $\Sigma_0$  represents the total number operator [1]. Moreover, we have the remarkable identify

$$\Sigma_1^2 + \Sigma_2^2 + \Sigma_3^2 = \Sigma_0^2 + 2\Sigma_0 \quad (\text{L.7})$$

which is a consequence of the commutation rules:

$$[\Sigma_0, \Sigma_j] = 0 \quad \forall j \in \{1, 2, 3\} \quad (\text{L.8a})$$

$$[\Sigma_1, \Sigma_2] = 2i\Sigma_3 \quad (\text{L.8b})$$

$$[\Sigma_2, \Sigma_3] = 2i\Sigma_1 \quad (\text{L.8c})$$

$$[\Sigma_3, \Sigma_1] = 2i\Sigma_2 \quad (\text{L.8d})$$

in sharp contrast to the classic Stokes parameters. We can say, therefore, that the noncommutability of the Stokes operators precludes the simultaneous measurement of the quantities represented by them.

The degree of polarization is constructed similarly as in the classic case:

$$P \equiv \frac{(\langle \Sigma_1 \rangle^2 + \langle \Sigma_2 \rangle^2 + \langle \Sigma_3 \rangle^2)^{1/2}}{\langle \Sigma_0 \rangle} \quad (\text{L.9})$$

The azimuth  $\psi$  and ellipticity  $\chi$  of the polarization ellipse are as given by Eqs. (3.1.155a–c):

$$\tan(2\psi) = \frac{\langle \Sigma_2 \rangle}{\langle \Sigma_1 \rangle} \quad (\text{L.10})$$

$$\tan(2\chi) = \frac{\langle \Sigma_3 \rangle}{(\langle \Sigma_1 \rangle^2 + \langle \Sigma_2 \rangle^2)^{1/2}} \quad (\text{L.11})$$

A physical understanding of this result can be gained from the consideration of a two-mode coherent state. In that case the degree of polarization is equal to 1; the coherent state corresponds to the classic completely polarized optical field [3].

It is further interesting to note that the noncommutability of the Stokes operators yields the following uncertainty principle:<sup>1</sup>

$$\langle(\Delta\Sigma_1)^2\rangle^{1/2}\langle(\Delta\Sigma_2)^2\rangle^{1/2} \geq |\langle\Sigma_3\rangle| \quad (\text{L.12})$$

Condition (L.12) may be derived directly from Eqs. (L.8b-d) and the Robertson relation for two Hermitian operators **A** and **B**

$$\Delta\mathbf{A}\Delta\mathbf{B} \geq \frac{1}{2}|\langle[\mathbf{A}, \mathbf{B}]\rangle| \quad (\text{L.13})$$

where  $\Delta\mathbf{A} = (\langle\mathbf{A}^2\rangle - \langle\mathbf{A}\rangle^2)^{1/2}$  is the variance of **A**. States for which an equality is achieved in inequality (L.13) are the minimum-uncertainty states.

The fact that Maxwell equations with dipole and current sources dependent linearly on the electromagnetic field are formally the same in quantum electrodynamics implies that the linear transformation between the incident and transmitted (or scattered) fields must also be valid in a quantum-mechanical picture. As we have seen above, the only proviso is that the classic field vector be treated as an operator. Following this way, one obtains the quantum-mechanical generalization of the Jones formalism. The operator that effects the transformation  $\mathbf{E}_i \rightarrow \mathbf{E}_o$  of the field operators is called the *Jones operator*.

## REFERENCES

1. J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons*, Addison-Wesley, Reading, MA, 1955. See also W. Heitler, *The Quantum Theory of Radiation*, Dover, New York, 1984.
2. W. H. Louisell, *Quantum Statistical Properties of Radiation*, Wiley, New York, 1973.
3. L. Mandel and E. Wolf, *Optical Coherence and Quantum Optics*, Cambridge Univ. Press, New York, 1995.
4. See the feature issue titled "Squeezed states of the electromagnetic field," *J. Opt. Soc. Am. B*4, 1987.

<sup>1</sup>For three observables **A**, **B**, **C** that obey a commutation relation  $[\mathbf{A}, \mathbf{B}] = i\mathbf{C}$ , it is easy to show that  $\langle\mathbf{A}^2\rangle\langle\mathbf{B}^2\rangle \geq \frac{1}{4}\langle\mathbf{C}\rangle^2$ .

## AUTHOR INDEX

---

- Abramovitz, 142  
Akkermans, 296  
Ampere, 41  
Anderson (P.), 294  
Anderson (D.G.M.), 231, 237  
Arago, 6  
Aravind, 133  
Asano, 259  
Azzam, 47, 207, 230, 248, 313, 317, 343  
  
Babinet, 7, 336  
Basov, 23  
Barakat, 22, 161, 167, 231, 237, 269, 372  
Barkla, 14  
Bartholinus, 3  
Bashara, 230, 248, 317  
Berek, 16  
Bernabeu, 231  
Berreman, 242  
Berry, 22, 132  
Bessel, 143, 154, 242, 353  
Bethe, 279, 382  
Bhandari, 135, 267  
Bicout, 240  
Biot, 6  
Bird, 24  
Birge, 19  
Blanc-Lapierre, 21  
Bloch, 123  
Bloembergen, 23, 25  
Bohren, 13, 369  
Born, 22  
Bottiger, 355  
Bourret, 22, 363  
Brewster, 5, 47, 247  
Brosseau, 167, 235, 240, 261, 269, 348, 372  
Bruhat, 14  
Bruscaglioni, 284  
Bunimovitch, 143  
  
Callies, 373  
Cauchy, 74, 201  
Chandrasekhar, 20, 252, 279  
  
Chipman, 377  
Chyba, 135  
Clark Jones, 17, 18, 75, 165, 179, 291, 305, 310, 313, 347, 373, 393  
Clausius, 260  
Collett, 247, 341  
Cotton, 11, 14, 337  
Coulomb, 52  
  
Debye, 13  
Descartes, 44, 249  
Deschamps, 20  
de Gennes, 324  
De Vito, 133  
Dirac, 74  
Dollfus, 23  
Drude, 12  
Dubridge, 19  
Dumontet, 21  
  
Elliott, 24  
Evans, 23  
  
Falkoff, 20  
Fano, 16  
Faraday, 9, 40, 243, 337  
Fercher, 345  
Fermat, 3  
Fizeau, 10  
Foucault, 7, 332  
Fourier, 72  
Franken, 25  
Fresnel, 5, 42, 244, 331, 336  
Freund, 292, 345  
Frobenius, 210, 374  
Fry, 355  
Fuller, 352  
  
Gabor, 19  
Gamo, 21  
Gantmacher, 366  
Gauss, 40  
Gerligand, 235

- Gil, 231  
 Givens, 217  
 Glan, 7, 332  
 Glauber, 24  
 Glazer, 336  
 Goldstone, 385  
 Goos, 19  
 Govi, 8  
 Graybill, 366  
 Green, 10, 382  
 Groosmuller, 16  
 Gu, 243  
 Gudkov, 373
- Haidinger, 7  
 Hale, 14  
 Hall, 24  
 Hanbury-Brown, 23, 116  
 Hänchen, 19  
 Hankel, 243  
 Helmhotz, 58, 260  
 Herapath, 7, 17  
 Herschel, 6  
 Hertz, 11, 334  
 Hilbert, 74, 128, 136  
 Hiltner, 23  
 Hopkins, 19  
 Hovenier, 256, 377  
 Hsü, 20  
 Huffman, 353  
 Huggins, 18, 317  
 Hurwitz, 19  
 Hunt, 353  
 Huyghens, 3  
 Huynen, 20  
 Hyde, 24
- Inoué, 24  
 Ising, 168
- Jasperson, 352  
 Jaswail, 159  
 Jauch, 24  
 Javan, 24  
 Jerrard, 20  
 Jones, *see Clark Jones*  
 Jordan, 16
- Kano, 100  
 Kemp, 352  
 Kenaugh, 20  
 Kerr, 10, 337
- Khintchine, 16, 98  
 Kim, 269  
 Kohlrausch, 10  
 Kostinski, 217  
 Kozlov, 137  
 Krein, 129  
 Krishnan, 18, 22, 252  
 Kronecker, 229  
 Kubota, 24  
 Kuhn, 10
- Lagendijk, 294  
 Lagrange, 172, 366  
 Lakeman, 16  
 Land, 17, 335  
 Langevin, 365  
 Langsdorf, 19  
 Lenard, 11  
 Leviero, 133  
 Lewis, 284  
 Lorentz, 11, 53, 62, 229, 260, 278  
 Lorenz, 12, 253, 260  
 Love, 182  
 Lu, 377  
 Luneberg, 19  
 Lyot, 19, 22
- MacDonald (J.E.), 20  
 MacDonald. (D.C.K.), 143  
 Mackintosh, 387  
 MacLaurin, 151, 313  
 MacNeillie, 332  
 Maiman, 24  
 Majorana, 13  
 Maker, 25  
 Malus, 4, 226  
 Mandel, 21, 115  
 Maret, 294, 387  
 Martinelli, 267  
 Mauguin, 14  
 Maxwell, 9, 37  
 Mehta, 100, 159  
 Michelson, 14, 37  
 Mie, 12, 13, 253, 278  
 Miller, 284  
 Millman, 129  
 Minkowski, 110  
 Mossotti, 260  
 Mouton, 11, 337  
 Mueller, 17, 18, 179, 210  
 Müller, 7
- Neumann, 5

- Newton (I.), 4  
Newton (R.G.), 253  
Nicol, 6, 332  
Noether, 51  
Nörrenberg, 7  
  
Oersted, 9  
Öhman, 22  
O'Neill, 165  
Onsager, 293  
Ore, 373  
  
Parke, 19, 181  
Parrish, 24  
Pancharatman, 22, 132, 164  
Parrent, 21  
Pasteur, 7  
Pauli, 106, 312  
Pellat-Finet, 20  
Pennypacker, 260  
Perrin, 17, 18, 22, 252, 256  
Planck, 15, 165, 372  
Pockels, 11, 337, 355  
Poincaré, 12, 123, 223  
Poisson, 10, 52  
Potekhin, 163  
Poynting, 41, 63, 95, 306  
Prigogine, 293  
Pritchard, 24  
Prokhorov, 23  
Prost, 324  
Purcell, 260  
  
Raab, 348  
Ramachandran, 20  
Ramaseshan, 20  
Ramsey, 237  
Rayleigh (Strutt), 13, 22, 140, 234, 253, 258, 277  
Reinitzer, 13, 324  
Riccati, 318  
Richartz, 20  
Rochon, 7, 332  
Roentgen, 14  
Rogers, 17  
Roman, 21  
Römer, 4  
Rohrlich, 24  
Rosen, 55, 373  
Rothen, 19  
Rozenberg, 20, 24  
  
Salpeter, 279, 382  
  
Samuel, 135  
Sato, 259  
Savage, 25  
Saxon, 22  
Schawlow, 23  
Schnatterly, 352  
Schuster, 14  
Schwarz, 201  
Schwarzschild, 14  
Sebeck, 7  
Senarmont, 7, 332  
Shurcliff, 19, 336  
Siebert, 387  
Sinclair, 20  
Snel, 44, 249  
Snyder, 182  
Solc, 23, 204  
Soleillet, 17  
Stamov, 20  
Stefan, 10  
Steeger, 345  
Stegun, 142  
Stokes, 7, 102, 113, 160, 179, 340, 396  
Sudarshan, 24  
Sylvester, 366  
  
Talbot, 6  
Terhune, 25  
Thomson, 355  
Thompson, 332  
Thomson (Lord Kelvin), 10  
Title, 336  
Tolhoek, 20  
Townes, 23  
Tyndall, 8  
Twiss, 23, 116  
Tykody, 293  
  
van Albada, 294  
van Cittert, 16  
van de Hulst, 20, 252, 256  
van der Mee, 377  
van Weert, 243  
Verdet, 10, 160, 196, 264  
Vernon, 18, 317  
von Helmholtz, 12  
von Laue, 14, 15, 165, 372  
von Neumann, 16, 166  
  
Waterman, 242  
Weber, 10  
Westfold, 20  
Whitney, 202

- Wien, 15, 165  
Wiener, (N.), 16, 21, 22, 98  
Wiener (O.), 11, 48  
Wolf (E.), 18, 21, 100, 111  
Wolf (P. E.), 294, 387  
Wollaston, 7, 331  
Wood, 19, 336  
  
Yariv, 207
- Yeh, 207, 243  
Young, (T.), 4  
Young (A.), 13  
  
Zapasski, 137  
Zeeman, 11  
Zernike, 16

## SUBJECT INDEX

---

- Absorbance, 190  
Absorber, 199  
Action-at-a-distance, 9  
Ampere law, 41  
Analytic signal, 19, 70  
Anisotropic medium, 18, 305  
Annihilation operator, 24, 393  
Arago–Fresnel interference laws, 6  
Autocorrelation function, 16,  
Autocorrelation tensor, 93  
Azimuth angle, 78, 319, 394
- Babinet compensator, 7, 337  
Backscattering enhancement, 293  
Backscattering matrix formalism, 21  
Beamsplitter, 334  
Berreman formalism, 242  
Bessel function, 242, 353  
Bethe-Salpeter equation, 279, 382  
Birefringence, 18, 197, 305, 317, 330  
Birefringent polarizer, 6  
Blackbody radiation, 22, 99, 363  
Bloch equation, 135  
Brewster angle, 5, 47, 247
- Calcite (Iceland spar), 3  
Cauchy formula, 196  
Cauchy–Schwarz inequality, 94, 108, 201  
Cavity, 64  
Cayley–Hamilton theorem, 119  
Characteristic function, 147  
Charge density, 39  
Chebyshev identity, 205  
Chiral asymmetry, 7  
Cholesteric phase, 324  
Circular birefringence, 5  
Coherence theory, 16, 21  
Coherency matrix formalism, 16, 21, 102, 185,  
    211  
Coherent state, 24  
Collett procedure, 341  
Commutation relation, 264, 393  
Compensator (retarder), 196, 217, 226, 336
- Completely polarized light, 7, 76, 187, 212  
Complex correlation factor, 94  
Complex polarization ratio representation, 20,  
    84, 206, 318  
Concavity property, 368  
Conductivity, 38  
Conformal bilinear transformation, 207  
Congruent transformation, 187, 211  
Convex set of states, 121, 128  
Corpuscular theory, 4  
Correlation function of analytic signal, 89  
Correlation length, 170  
Cotton–Mouton effect, 11  
Coulomb gauge, 52  
Coupled dipole approximation, 260  
Covariance matrix of instantaneous Stokes  
    parameters, 116  
Creation opeator, 24, 393  
Cross-correlation function, 16  
Cross covariance, 93  
Cross-spectral correlation tensor, 97  
Cross spectral purity, 22  
Cumulant, 151  
Current density, 38  
 $C_{xy}$  symmetry, 256  
 $C_{zy}$  symmetry, 256
- Degree of coherence, 16, 108  
Degree of polarization, 111, 323, 368, 376, 394  
Density matrix formalism, 105, 368, 393  
Depolarization, 235, 241  
Depolarization length scale, 276  
Depolarizer, 222, 232, 283  
Descartes–Snel law, 44  
Destructive interference, 4  
Deterministic optical device, 179  
Diattenuation, 186, 202, 208  
Dichroism, 10, 305, 330  
Differential Jones matrix, 309, 312, 317  
Differential Mueller matrix, 313, 325  
Differential polarization matrices, 309  
Differential polarized spectroscopy, 348  
Diffusion constant, 276

- Diffusion wave spectroscopy, 283, 387  
 Dirac delta function, 90  
 Direct (Kronecker) product of matrices, 105, 229  
 Dispersion relation, 60  
 Displacement current, 10  
 Division-of-amplitude photopolarimetry, 343  
 Double-split optical interference experiment, 4  
 Double refraction, 3  
 Duality transformation, 55
- Eccentricity, 70  
 Eigenvalue-eigenvector decomposition, 185  
 Einstein's special theory of relativity, 11  
 Elastic mean free path, 276  
 Elastic scattering, 18  
 Electric displacement, 38  
 Electric field vector, 15, 38  
 Electromagnetic theory, 8  
 Ellipse of polarization, 70, 317  
 Ellipsometry, 19, 248  
 Elliptically polarized, 11  
 Elliptic integral, 142  
 Ellipticity angle, 76, 78, 319, 394  
 Elliptic polarization, 81, 107  
 Emisionnist, 4  
 Enantiomorphism, 7  
 Entropy of radiation, 15, 164, 372  
 Entropy production, 209, 281, 289  
 Equation of continuity, 39  
 Ergodic random process, 71  
 Ethereal matter, 3, 5  
 Expectation value, 107  
 Extraordinary ray, 3
- Fabry–Perot interferometer, 23  
 Faraday effect, 9  
 Faraday law of induction, 40  
 Fast axis, 196  
 Fork concept, 21  
 Fourier series, 353  
 Fourier transform, 16, 71  
 Fourth order effect, 3  
 Fresnel equations, 5, 12, 42, 244  
 Fresnel equation of wave normals, 308  
 Fresnel rhomb, 336  
 Frobenius norm, 201, 374
- Gain (transmittance), 187, 212, 231, 271  
 Gauge transformation, 51  
 Gaussian random process, 4, 117, 139, 270  
 Gaussian spectrum, 73  
 Gauss' theorem, 40, 64  
 Gell–Mann matrices, 119
- Geodesic triangle, 133  
 Geometric phase, 22, 131  
 Generalized Stokes parameters, 118, 371  
 Glan–Foucault polarizer, 7, 332  
 Glan–Thompson polarizer, 332, 343  
 Goos–Hänchen effect, 19  
 Green's function, 280, 285, 382, 387  
 Group velocity, 60  
 Gyration vector, 38
- Haidinger's brush, 7, 19  
 Half-waveplate, 198  
 Handedness, 78  
 Hankel function, 243  
 Harmonic oscillator, 64  
 Harmonic solution, 54, 58  
 Helicity basis, 75  
 Helmholtz equation, 57  
 Helmholtz free energy, 170  
 Herapathite, 7  
 Hermitian conjugate, 105  
 Hermitian matrix, 191  
 Hilbert space, 128, 136  
 Hilbert transform, 74  
 Holography, 20  
 Homogeneous material medium, 39, 57  
 Homogeneous plane wave, 58  
 Hypergeometric function, 141
- Improper function, 74  
 Incoherent mixture of partially polarized light beams, 368  
 Inhomogeneous plane wave, 58  
 Instantaneous ellipse for the electric field, 76  
 Instantaneous envelope, 76  
 Instantaneous phase, 76  
 Instantaneous Stokes parameters, 102  
 Instantaneous Stokes vector, 102  
 Integration response time, 71  
 Intensity (irradiance), 78  
 Interference filter, 23  
 Internal energy, 170  
 Invariance transformation, 51  
 Irreversibility, 291  
 Ising model, 169  
 Isotropic medium, 57  
 Isotropic suspensions of particles, 18, 378  
 Isotropic turbulence of an incompressible fluid, 363
- Joint bivariate probability density function, 139  
 Jones formalism, 18  
 Jones matrix, 180, 269, 346  
 Jones operator, 395

- Jones procedure, 347  
 Jones vector, 75, 179  
 Jordan canonical form, 204  
 Jordan lemma, 148
- Kaleidoscope, 5  
 Kerr effect, 10  
 Krein–Millman theorem, 129  
 Kronecker product of matrices, 229  
 Kurtosis, 151
- Ladder approximation, 383  
 Lagrange multipliers, 172  
 Lagrange–Sylvester theorem, 366  
 Lamellae formalism, 315  
 Langevin function, 365  
 Laser, 23  
 Left-handed polarization, 78  
 Left-handed vector, 75  
 Line of force, 9  
 Linearly polarized light, 11  
 Linear momentum flux, 41  
 Liquid crystals, 13, 324  
 Lissajous diagram, 82  
 Localization, 294  
 Lorentz gauge, 53  
 Lorentz lemma, 62  
 Lorentz spectrum, 73, 159  
 Lorentz transformation, 229  
 Lyot depolarizer, 19
- MacDonald–Bunimovitch relation, 143  
 MacLaurin expansion, 151, 313  
 MacNeill prism, 334  
 Magnetic field vector, 38  
 Magnetic induction, 38  
 Magnetogyration coefficient, 38  
 Major axis of the ellipse, 77  
 Maser, 23  
 Material equations, 38  
 Maximum entropy principle, 171, 378  
 Maximum gain, 188, 212, 272  
 Maxwell equations, 10, 37, 382  
 Maxwell theory, 9  
 Minimum entropy production, 293  
 Minimum gain, 188, 212, 272  
 Minkowski metric matrix, 110  
 Minor axis of the ellipse, 77  
 Mirror, 200, 222  
 Mixed state, 110, 126, 129  
 Modified Bessel function, 140, 143, 145, 154  
 Monochromatic radiation, 20, 57, 69  
 Monte Carlo simulation, 277, 283  
 Mueller formalism, 18, 351
- Mueller–Jones matrix, 228, 236, 273, 314  
 Mueller matrix, 24, 210, 245  
 Mueller matrix imaging polarimeter, 356  
 Multilayered medium, 203  
 Multiple scattering, 240, 275, 383, 387
- Natural light, 8  
 Negative exponential law, 149  
 Nematic phase, 324  
 Nicol polarizer, 6, 332  
 Noether theorem, 51  
 Nonimage forming optical instrument, 18  
 Nonlinear optics, 25  
 Nonplane wave, 22  
 Normalized instantaneous Stokes parameters, 153  
 Normalized Jones vector, 75  
 Normalized Stokes parameters, 119  
 Normal mode, 63  
 Nörrenberg reflecting polariscope, 7
- $O(3)$ , 191  
 Observable, 21, 393  
 Optical activity, 5, 6, 30, 196  
 Optical depth, 284  
 Optical equivalence, 113  
 Optical fiber, 329  
 Optically induced birefringence, 25  
 Optical microscope, 6  
 Optical rotation, 10  
 Optical symmetry, 306  
 Optimal polarization null concept, 21  
 Ordinary ray, 3  
 Orthogonalization, 207  
 Orthonormal pair of vectors, 84, 189  
 Overdetermination, 356  
 Overgain, 215  
 Overpolarization, 214
- Pancharatnam rule, 132  
 Paraxial approximation, 18  
 Partially polarized light, 7, 70  
 Passive optical medium, 179, 200, 212  
 Pauli matrices, 105  
 Pencil of radiation, 15  
 Permeability, 38  
 Permittivity, 38, 306  
 Phase angle, 77  
 Phase conjugation mirror, 185, 262  
 Phase contrast microscope, 16  
 Phase velocity, 60, 308  
 Photoelasticity (strain birefringence), 5  
 Photon polarization, 20, 392  
 Physical realizability constraint, 200, 213

- Pile-of-plates configuration, 6, 247, 332  
 Pince à tourmalines, 6  
 Pitch of the helical structure, 326  
 Planar interface, 5  
 Plane of constant amplitude, 59  
 Plane of constant phase, 58  
 Plane of vibration, 58  
 Plane wave, 57  
 Pockels' effect, 11  
 Poincaré sphere, 12, 22, 123, 130, 223, 241, 263, 376  
 Poincaré vector, 124, 136  
 Pointlike suspension, 277  
 Polar decomposition, 201  
 Polariscope, 6  
 Polarizance, 208  
 Polarization, 208  
     by reflection, 5, 331  
     by selective absorption, 334  
     by transmission, 332  
 Polarization cycle, 131  
 Polarization entropy, 208, 227, 366  
 Polarization matrices, 17  
 Polarization-modulated nephelometer, 353  
 Polarization modulation method, 352  
 Polarization radar, 21  
 Polarization transfer, 233  
 Polarizer, 192, 220, 225, 331  
 Polarizing angle, 5  
 Polarizing microscope, 6, 14, 24  
 Polaroid, 335  
 Polychromatic light, 22, 164  
 Polystyrene sphere, 286  
 Power spectrum, 16, 69  
 Poynting theorem, 62  
 Poynting vector, 41, 61, 95  
 Principal value in Cauchy's sense, 74  
 Probability density function, 91, 139, 147, 282, 345, 363  
 Projection matrix, 193  
 Projector, 110  
 Pure state, 110, 126, 129  
  
 Quantum optics, 24  
 Quantum theory, 15  
 Quarter-waveplate, 197, 336  
 Quartz, 305  
 Quasimonochromatic, 16, 69  
 Quaternions, 20  
  
 Radiation fluctuations, 15  
 Radiative transfer, 20, 279  
 Random medium, 20, 179, 257, 268, 293, 324  
 Random process, 15  
  
 Random walk, 163  
 Rank, 128, 191, 374  
 Rayleigh distribution, 140  
 Rayleigh scattering, 13, 258, 277  
 Reciprocity, 8, 12, 61, 260  
 Reduction property, 115  
 Reflection coefficient, 45  
 Reflectivity, 45  
 Refractive index, 5, 44  
 Residues, 148  
 Retardation, 186, 202  
 Retarded potential solution, 53  
 Riccati equation 318  
 Right-handed polarization, 78  
 Right-handed vector, 75  
 Rochon prism, 7  
 Rotator, 195, 219, 332  
 Round-trip, 266  
 Runge-Kutta integration scheme, 327  
  
 Scalar invariant, 119, 174  
 Scalar potential, 51  
 Scattering amplitude matrix, 254  
 Scattering of light by particles, 8, 253  
 Scattering medium, 18  
 Second-order correlation 16, 96  
 Second harmonics, 25  
 Senarmont compensator, 7, 332  
 Serial device, 203, 211  
 Sheet polarizer, 17  
 SI system of units, 38  
 Siegert-Reed theorem, 387  
 Similarity invariant, 190, 208  
 Similarity transformation, 184  
 Single scattering, 253  
 Singular value, 189, 374  
 Size parameter, 276  
 Skewness factor, 151  
 Slow axis, 196  
 Smectic phase, 324  
 SO(3), 128, 225, 241  
 Solc filter, 204  
 Spatial coherence of light, 19  
 Spatial fluctuation 69  
 Spatial inversion, 55  
 Specific heat, 170  
 Speckle, 268, 292, 345  
 Spectral coherency matrix, 22, 109  
 Spectral decomposition theorem, 109, 117, 366  
 Spectral density matrix, 109  
 Spectral purity, 115  
 Speed of light, 38  
 Spherical triangle, 134  
 Spherical wave, 4

- Spheroidal particle, 259  
 Spin echo, 137  
 Spin- $\frac{1}{2}$  particle, 108  
 Spin-1 particle, 392  
 Spinor, 17  
 Square-integrable random process, 71  
 State of polarization, 70  
 Stationary random process, 71  
 Stereographic projection, 12, 123, 130, 263  
 Stokes operators, 24, 393  
 Stokes parameters, 7, 18, 102, 320, 339  
 Stokes' procedure, 340  
 Stokes vector, 104, 210, 313  
 Stokes vector space, 124  
 Stokes–Verdet–Barakat conditions, 160  
 Stream of particles, 4  
 $SU(2)$ , 107, 111, 130, 170, 191  
 $SU(N)$ , 119  
 Surface density of bound charge, 42  
 Surface density of current, 42  
 Symmetry transformations, 252
- Tangent law, 5  
 Temperature of polarization, 168  
 Temporal electric field autocorrelation function, 387  
 Temporal fluctuation, 69  
 Temporal intensity autocorrelation function, 387  
 Thermal radiation, 24, 69  
 Thermodynamics of radiation, 15, 168, 291  
 Time-averaged Poynting vector, 60  
 Time of coherence, 70  
 Time-integrated Stokes parameter, 158  
 Time reversal, 51, 54, 261, 294  
 T-matrix method, 242  
 Total internal reflection, 46
- Tourmaline, 305  
 Trace condition, 231, 237, 275  
 Transmission coefficient, 45  
 Transmissivity, 45  
 Transport mean free path, 276, 285  
 Transverse nature of light, 4, 11  
 Two-level system, 107, 135
- Unitary matrix, 75, 190, 202  
 Unitary transformation, 191  
 Unpolarized light, 7, 10, 76, 160, 170, 186, 212
- van Cittert–Zernike theorem, 16  
 Vector potential, 51, 63  
 Velocity of light, 4, 60  
 Verdet constant, 10, 196, 264  
 Videopolarimetry, 345  
 Visibility of interference fringes, 16, 190  
 Volume charge density, 38  
 Volume density of energy, 41, 61  
 von Neumann entropy, 166
- Wavefront, 4, 59  
 Wave nature of light, 4  
 Wave vector, 58  
 White light, 16  
 White noise, 276  
 Wiener–Khintchine theorem, 16, 97  
 Wiener experiment, 11, 49  
 Wire-grid polarizer, 11, 24, 334  
 Wollaston prism, 7, 333
- X ray, 11, 14
- Young experiments, 4
- Zeeman's effect, 11

Lightning Source UK Ltd.  
Milton Keynes UK  
UKOW022020111011

180094UK00005B/24/A



9 780471 143024