SPIE PRESS | Tutorial Text

Polarization of Light with Applications in Optical Fibers

Arun Kumar
Ajoy Ghatak

Polarization
of Light with Applications in
Optical Fibers

Tutorial Texts Series

- *Modeling the Imaging Chain of Digital Cameras*, Robert D. Fiete, Vol. TT92
- *Bioluminescence and Fluorescence for* In Vivo *Imaging*, Lubov Brovko, Vol. TT91
- *Polarization of Light with Applications in Optical Fibers*, Arun Kumar, Ajoy Ghatak, Vol. TT90
- *Digital Fourier Optics: A MATLAB Tutorial*, David G. Voeltz, Vol. TT89
- *Optical Design of Microscopes*, George Seward, Vol. TT88
- *Analysis and Evaluation of Sampled Imaging Systems*, Richard H. Vollmerhausen, Donald A. Reago, Ronald Driggers, Vol. TT87
- *Nanotechnology: A Crash Course*, Raúl J. Martin-Palma and Akhlesh Lakhtakia, Vol. TT86
- *Direct Detection LADAR Systems*, Richard Richmond, Stephen Cain, Vol. TT85
- *Optical Design: Applying the Fundamentals*, Max J. Riedl, Vol. TT84
- *Infrared Optics and Zoom Lenses, Second Edition*, Allen Mann, Vol. TT83
- *Optical Engineering Fundamentals, Second Edition*, Bruce H. Walker, Vol. TT82
- *Fundamentals of Polarimetric Remote Sensing*, John Schott, Vol. TT81
- *The Design of Plastic Optical Systems*, Michael P. Schaub, Vol. TT80
- *Fundamentals of Photonics*, Chandra Roychoudhuri, Vol. TT79
- *Radiation Thermometry: Fundamentals and Applications in the Petrochemical Industry*, Peter Saunders, Vol. TT78
- *Matrix Methods for Optical Layout*, Gerhard Kloos, Vol. TT77
- *Fundamentals of Infrared Detector Materials*, Michael A. Kinch, Vol. TT76
- *Practical Applications of Infrared Thermal Sensing and Imaging Equipment, Third Edition*, Herbert Kaplan, Vol. TT75
- *Bioluminescence for Food and Environmental Microbiological Safety*, Lubov Brovko, Vol. TT74
- *Introduction to Image Stabilization*, Scott W. Teare, Sergio R. Restaino, Vol. TT73
- *Logic-based Nonlinear Image Processing*, Stephen Marshall, Vol. TT72
- *The Physics and Engineering of Solid State Lasers*, Yehoshua Kalisky, Vol. TT71
- *Thermal Infrared Characterization of Ground Targets and Backgrounds, Second Edition*, Pieter A. Jacobs, Vol. TT70
- *Introduction to Confocal Fluorescence Microscopy*, Michiel Müller, Vol. TT69
- *Artificial Neural Networks: An Introduction*, Kevin L. Priddy and Paul E. Keller, Vol. TT68
- *Basics of Code Division Multiple Access (CDMA)*, Raghuveer Rao and Sohail Dianat, Vol. TT67
- *Optical Imaging in Projection Microlithography*, Alfred Kwok-Kit Wong, Vol. TT66
- *Metrics for High-Quality Specular Surfaces*, Lionel R. Baker, Vol. TT65
- *Field Mathematics for Electromagnetics, Photonics, and Materials Science*, Bernard Maxum, Vol. TT64
- *High-Fidelity Medical Imaging Displays*, Aldo Badano, Michael J. Flynn, and Jerzy Kanicki, Vol. TT63
- *Diffractive Optics–Design, Fabrication, and Test*, Donald C. O'Shea, Thomas J. Suleski, Alan D. Kathman, and Dennis W. Prather, Vol. TT62
- *Fourier-Transform Spectroscopy Instrumentation Engineering*, Vidi Saptari, Vol. TT61
- *The Power- and Energy-Handling Capability of Optical Materials, Components, and Systems*, Roger M. Wood, Vol. TT60
- *Hands-on Morphological Image Processing*, Edward R. Dougherty, Roberto A. Lotufo, Vol. TT59
- *Integrated Optomechanical Analysis*, Keith B. Doyle, Victor L. Genberg, Gregory J. Michels, Vol. TT58
- *Thin-Film Design: Modulated Thickness and Other Stopband Design Methods*, Bruce Perilloux, Vol. TT57
- *Optische Grundlagen für Infrarotsysteme*, Max J. Riedl, Vol. TT56
- *An Engineering Introduction to Biotechnology*, J. Patrick Fitch, Vol. TT55
- *Image Performance in CRT Displays*, Kenneth Compton, Vol. TT54
- *Introduction to Laser Diode-Pumped Solid State Lasers*, Richard Scheps, Vol. TT53
- *Modulation Transfer Function in Optical and Electro-Optical Systems*, Glenn D. Boreman, Vol. TT52

(For a complete list of Tutorial Texts, see http://spie.org/x651.xml.)

Polarization
of Light with Applications in
Optical Fibers

Arun Kumar
Ajoy Ghatak

Tutorial Texts in Optical Engineering
Volume TT90

SPIE
PRESS
Bellingham, Washington USA

Library of Congress Cataloging-in-Publication Data

Kumar, Arun, 1952-
Polarization of light with applications in optical fibers / Arun Kumar, Ajoy Ghatak.
p. cm. -- (Tutorial texts in optical engineering ; v. TT90)
Includes bibliographical references and index.
ISBN 978-0-8194-8215-0
1. Fiber optics--Mathematics. 2. Optical fibers--Mathematical models. 3. Polarization (Light)--Mathematical models. I. Ghatak, A. K. (Ajoy K.), 1939- II. Title.
TA1800.K86 2010
621.36'92--dc22

2010038731

Published by

SPIE
P.O. Box 10
Bellingham, Washington 98227-0010 USA
Phone: +1 360.676.3290
Fax: +1 360.647.1445
Email: books@spie.org
Web: http://spie.org

Printed in the United States of America.

Introduction to the Series

Since its inception in 1989, the Tutorial Texts (TT) series has grown to cover many diverse fields of science and engineering. The initial idea for the series was to make material presented in SPIE short courses available to those who could not attend and to provide a reference text for those who could. Thus, many of the texts in this series are generated by augmenting course notes with descriptive text that further illuminates the subject. In this way, the TT becomes an excellent stand-alone reference that finds a much wider audience than only short course attendees.

Tutorial Texts have grown in popularity and in the scope of material covered since 1989. They no longer necessarily stem from short courses; rather, they are often generated independently by experts in the field. They are popular because they provide a ready reference to those wishing to learn about emerging technologies or the latest information within their field. The topics within the series have grown from the initial areas of geometrical optics, optical detectors, and image processing to include the emerging fields of nanotechnology, biomedical optics, fiber optics, and laser technologies. Authors contributing to the TT series are instructed to provide introductory material so that those new to the field may use the book as a starting point to get a basic grasp of the material. It is hoped that some readers may develop sufficient interest to take a short course by the author or pursue further research in more advanced books to delve deeper into the subject.

The books in this series are distinguished from other technical monographs and textbooks in the way in which the material is presented. In keeping with the tutorial nature of the series, there is an emphasis on the use of graphical and illustrative material to better elucidate basic and advanced concepts. There is also heavy use of tabular reference data and numerous examples to further explain the concepts presented. The publishing time for the books is kept to a minimum so that the books will be as timely and up-to-date as possible. Furthermore, these introductory books are competitively priced compared to more traditional books on the same subject.

When a proposal for a text is received, each proposal is evaluated to determine the relevance of the proposed topic. This initial reviewing process has been very helpful to authors in identifying, early in the writing process, the need for additional material or other changes in approach that would serve to strengthen the text. Once a manuscript is completed, it is peer reviewed to ensure that chapters communicate accurately the essential ingredients of the science and technologies under discussion.

It is my goal to maintain the style and quality of books in the series and to further expand the topic areas to include new emerging fields as they become of interest to our reading audience.

James A. Harrington
Rutgers University

Contents

Preface

The polarization phenomena associated with light waves are extremely important in many areas of photonics. Many important polarization-based devices have been developed, including Faraday rotators; in-line fiber optic components, such as polarizers, wave plates, and polarization controllers; and sensors, such as fiber optic current sensors, and fiber optic gyroscopes. In order to understand the workings of such photonic devices and to improve their design, sound knowledge of the basic concepts involved in polarization is required.

Furthermore, in optical communication systems, polarization mode dispersion has become an extremely important issue, particularly for very high-bit-rate (>10 Gb) systems. Polarization mode dispersion arises because of random birefringence present in a practical optical fiber. The birefringence that causes polarization mode dispersion in optical fibers may be linear, circular, or, in general, elliptical. In order to understand the nature of polarization mode dispersion and to control or reduce it, one must know how the various types of birefringent media affect the polarization state of the guided light while it propagates through an optical fiber. Thus, it has become almost essential for most engineers (working in the general area of photonics) to have a basic knowledge of the polarization phenomena and associated concepts, as well as the basic methods of analysis, such as Jones matrices, Stokes parameters, and the Poincaré sphere.

In this book, our aim is to provide in one source all of the basic concepts and methods involved in the understanding and design of various photonic devices, keeping the discussions as succinct as possible. The Poincaré sphere representation of polarized light is a very important method that is not discussed in sufficient detail in most of the literature. Therefore, we have devoted an entire chapter to Poincaré sphere representation, including several numerical examples to make the method very clear. This book works through all steps using many examples; therefore, even undergraduate students should be able to follow along without much difficulty.

We have been teaching various aspects of polarization to our undergraduate students and to our master's students at the Indian Institute of Technology Delhi (IITD). This book has grown out of the lecture notes that we have prepared over the last 25 years. We have also used this material in several short courses organized at IITD and at other institutions.

We thank our colleagues in the physics department of IITD for many helpful discussions—in particular, we thank Profs. B. D. Gupta, B. P. Pal, A. Sharma, M. R. Shenoy, and K. Thyagarajan, and Dr. Ravi Varshney for research collab-

oration and useful discussions that have helped us in improving the presentation. We are also thankful to our research students Ms. Triranjita Srivastava and Mr. Saurabh Mani Tripathi, who helped us at various stages through their suggestions, in carrying out some of the calculations, and in creating some of the diagrams. Finally, we are grateful to our wives, Shobhita and Gopa, for their patience and understanding.

We dedicate this book to the memory of our parents.

Arun Kumar
Ajoy Ghatak
New Delhi
January 2011

Acronyms and Abbreviations

ADP	ammonium dihydrogen phosphate
CP	circularly polarized
CT	cross talk
DGD	differential group delay
DOP	degree of polarization
DSF	dispersion-shifted fiber
DUT	device under test
EM	electromagnetic
EP	elliptically polarized
Hi-Bi	high-birefringence (fiber)
HWP	half-wave plate
JME	Jones matrix eigenanalysis
KDP	potassium dihydrogen phosphate
LA	linear analyzer
LCP	left-circularly polarized
LD	laser diode
LED	light-emitting diode
LEP	left-elliptically polarized
LHP	linear horizontal polarizer
LP	linearly polarized
LVP	linear vertical polarizer
MFD	mode field diameter
MMM	Mueller matrix method
NZDSF	nonzero-dispersion-shifted fiber
OTDR	optical time-domain reflectometer
PANDA	polarizing-maintaining and absorption-reducing (fiber)
PC	polarization controller
PMD	polarization mode dispersion
PMF	polarization-maintaining fiber
PS	Poincaré sphere
PSA	Poincaré sphere analysis
PSP	principal states of polarization
QWP	quarter-wave plate
RCP	right-circularly polarized
REP	right-elliptically polarized
rms	root-mean-square

SAP	stress-applying part
SIRF	spin-induced reduction factor
SMF	single-mode fiber
SOP	state of polarization
SP	surface plasmon
SPP	surface plasmon polariton
SPSM	single-polarization single-mode (fiber)
SSMF	standard single-mode fiber
TE	transverse electric
TGG	terbium gallium garnet
TM	transverse magnetic
TMEF	two-mode elliptic-core fiber
TMF	two-mode fiber
UP	unpolarized
YIG	yttrium iron garnet

Chapter 1
Introduction

A solid understanding of the polarization characteristics of light waves has now become essential in the use of numerous optical devices in a variety of applications. Optical devices may possess linear, circular, or, in general, elliptical birefringence. While linear birefringence is quite simple to analyze, circular and elliptical birefringencies are more difficult to work with. Jones matrix and Poincaré sphere approaches are two powerful techniques used in polarization-based applications. In optical communication systems, polarization mode dispersion (PMD) is a serious issue for ultrahigh-bit-rate (>10 Gb) data transmission.

A large number of important fiber optic sensors and components have recently been developed, including fiber optic gyroscopes, current sensors, and in-line fiber optic polarization-based components such as polarizers, wave plates, and polarization controllers, all of which use polarization effects in optical fibers. Accordingly, sound knowledge of the polarization characteristics of light waves has become almost essential for all working engineers. This book aims to develop various concepts from first principles so that a working engineer is able to not only understand the basic physics of polarization-based devices, but also to improve on the design and characteristics of these devices.

All electromagnetic phenomena can be said to follow from Maxwell's equations. In Chapter 2, we write these equations and their constitutive relations. We use Maxwell's equations to derive the wave equation and study the consequences. In Chapter 3, we start with the plane wave solutions of Maxwell's equations and discuss the various states of polarization—namely, linear polarization, circular polarization, and elliptical polarization. We discuss methods of producing such polarization states as well as experimental configurations used to analyze them.

When an unpolarized light beam enters an anisotropic medium (such as calcite), it usually splits into two linearly polarized beams. These two linearly polarized waves, in general, propagate with different phase and ray velocities and hence see different refractive indices. This phenomenon is known as *double refraction*, or *birefringence*, and is of great importance as it is used to realize many polarization-based devices. In Chapter 4, we study plane wave propagation in an anisotropic medium and discuss some applications.

In various applications of polarized light, one often needs to know how the polarization of the incident wave changes when it is passed through a polarizer or a birefringent medium. This can be described by using either Jones calculus or the Poincaré sphere. In Chapter 5, we discuss the Jones calculus approach, which involves the representation of a given state of polarization (SOP) by a two-component vector and uses 2×1 matrices to calculate the effect of a polarizer or a birefringent medium on a given SOP.

The Jones vector representation of polarized light (discussed in Chapter 5) is quite simple and straightforward. However, this representation describes only polarized light, not partially polarized light. In Chapter 6, we discuss a more versatile representation of the polarization state of any light wave, which can be completely polarized, unpolarized, or partially polarized. In Chapter 7, we discuss the Poincaré sphere representation, which is a simple and extremely useful geometrical representation of various polarization states and their evolution through a birefringent medium. We show that this representation makes the formulation (as well as the explanation of a difficult problem) very easy due to the graphical depiction of the actions of the polarizers and birefringent media involved.

In Chapter 8, we study the propagation and polarization characteristics of single-mode fibers, which are widely used in most fiber-based devices. In Chapter 9, we discuss the origins and applications of birefringence in optical fibers. As mentioned earlier, PMD is a serious issue for ultrahigh-bit-rate (>10 Gb) fiber optic communication systems; therefore, in Chapter 10, we discuss PMD.

Chapter 2
Maxwell's Equations and Plane Wave Solutions

2.1 Introduction

All electromagnetic phenomena can be said to follow from Maxwell's equations. In Section 2.2, we will write these equations and the constitutive relations and derive the 3D wave equation in a dielectric. In Section 2.3, we will show that plane waves satisfy Maxwell's equations, and will study the properties of plane waves. Finally, in Section 2.4, we will give a physical interpretation of the Poynting vector.

2.2 Maxwell's Equations and the Wave Equation in an Isotropic Dielectric

Maxwell's equations are based on experimental observations and are given by the following equations:

$$\text{div } \mathbf{D} \ (= \nabla \cdot \mathbf{D}) = \rho, \tag{2.1}$$

$$\text{div } \mathbf{B} \ (= \nabla \cdot \mathbf{B}) = 0, \tag{2.2}$$

$$\text{curl } \mathbf{E} \ (= \nabla \times \mathbf{E}) = -\frac{\partial \mathbf{B}}{\partial t}, \tag{2.3}$$

and

$$\text{curl } \mathbf{H} \ (= \nabla \times \mathbf{H}) = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}, \tag{2.4}$$

where ρ represents the charge density, **J** the current density, and **E, D, B**, and **H** represent the electric field, electric displacement, magnetic induction, and magnetic field, respectively. The preceding equations can be solved only if the *constitutive relations* that relate **D** to **E**, **B** to **H**, and **J** to **E** are known. We will

assume the medium to be isotropic such that **D** is in the same direction as **E**. In this case, the constitutive relations are given by

$$\mathbf{D} = \varepsilon\, \mathbf{E}, \tag{2.5}$$

$$\mathbf{B} = \mu\, \mathbf{H}, \tag{2.6}$$

and

$$\mathbf{J} = \sigma\, \mathbf{E}, \tag{2.7}$$

where ε, μ, and σ are known respectively as the dielectric permittivity, magnetic permeability, and conductivity of the medium. In an anisotropic medium (such as calcite or quartz), the displacement vector **D**, in general, is not in the direction of the electric field **E**; we will discuss propagation in such anisotropic media in Chapter 4. Further, for a charge-free dielectric, we have $\rho = 0$ and $\sigma = 0$; hence,

$$\mathbf{J} = 0. \tag{2.8}$$

Using the relations given in Eqs. (2.5)–(2.7), Maxwell's equations simplify to

$$\nabla \cdot \mathbf{D} = 0, \tag{2.9}$$

$$\nabla \cdot \mathbf{H} = 0, \tag{2.10}$$

$$\nabla \times \mathbf{E} = -\mu\, \frac{\partial \mathbf{H}}{\partial t}, \tag{2.11}$$

and

$$\nabla \times \mathbf{H} = \varepsilon\, \frac{\partial \mathbf{E}}{\partial t}. \tag{2.12}$$

If we take the curl of Eq. (2.11), we obtain

$$\nabla \times (\nabla \times \mathbf{E}) = -\mu\, \frac{\partial}{\partial t} \nabla \times \mathbf{H} = -\varepsilon \mu\, \frac{\partial^2 \mathbf{E}}{\partial t^2}, \tag{2.13}$$

where we used Eq. (2.12) to substitute **H** in terms of **E**. If we now use the vector identity

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla\, (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}, \tag{2.14}$$

we obtain

$$\nabla (\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\varepsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{2.15}$$

In Eq. (2.14), the operator ∇^2 on vector $\mathbf{E}$ is defined as

$$\nabla^2 \mathbf{E} = \hat{\mathbf{x}} \nabla^2 E_x + \hat{\mathbf{y}} \nabla^2 E_y + \hat{\mathbf{z}} \nabla^2 E_z, \tag{2.16}$$

where

$$\nabla^2 E_x = \text{div grad } E_x. \tag{2.17}$$

Now,

$$\nabla \cdot \mathbf{D} = (\nabla \varepsilon) \cdot \mathbf{E} + \varepsilon (\nabla \cdot \mathbf{E}). \tag{2.18}$$

For a homogeneous medium, Eq. (2.9) reduces to

$$\nabla \cdot \mathbf{E} = 0; \tag{2.19}$$

thus, Eq. (2.15) becomes

$$\nabla^2 \mathbf{E} = \varepsilon \mu \frac{\partial^2 \mathbf{E}}{\partial t^2}. \tag{2.20}$$

Similarly, taking the curl of Eq. (2.12), one can show that **H** satisfies the following equation:

$$\nabla^2 \mathbf{H} = \varepsilon \mu \frac{\partial^2 \mathbf{H}}{\partial t^2}. \tag{2.21}$$

Equations (2.20) and (2.21) are known as the 3D wave equations and show that each Cartesian component of **E** and **H** satisfies the scalar wave equation

$$\nabla^2 \psi = \frac{1}{v^2} \frac{\partial^2 \psi}{\partial t^2}, \tag{2.22}$$

where

$$v = \frac{1}{\sqrt{\varepsilon \mu}}. \tag{2.23}$$

Equation (2.22) is known as the wave equation because whenever, from physical considerations, we are able to derive such an equation, we can predict the existence of waves whose velocity will be given by v.

In free space,

$$\varepsilon = \varepsilon_0 = 8.8542 \times 10^{-12}\ \mathrm{C^2/N\,m^{-2}}, \tag{2.24}$$

$$\mu = \mu_0 = 4\pi \times 10^{-7}\ \mathrm{N\,s^2 / C^2}, \tag{2.25}$$

and the speed of propagation of an electromagnetic wave is given by

$$v = c = \frac{1}{\sqrt{8.8542 \times 10^{-12} \times 4\pi \times 10^{-7}}} = 2.99794 \times 10^8\ \mathrm{m/s}. \tag{2.26}$$

Around 1860, Maxwell (using the laws of electricity and magnetism) derived the wave equation, predicted the existence of electromagnetic waves, and calculated the speed of these waves to be about 3×10^8 m/s; this he found to be very close to the velocity of light as measured by Fizeau in 1849. Based only on the closeness of these two numbers and with "faith in the rationality of nature," he proposed the electromagnetic theory of light and predicted that light must be an electromagnetic wave. [We should also mention here that the physical laws described by Eqs. (2.1)–(2.4) were known before Maxwell; he introduced only the term $\partial D/\partial t$ (which is the concept of displacement current) in Eq. (2.4); it was the presence of this term that allowed him to derive the wave equation and to predict the existence of electromagnetic waves.] It was not until 1888 that Heinrich Hertz carried out experiments that could produce and detect electromagnetic waves of frequencies smaller than those of light. Hertz showed that the velocity of electromagnetic waves that he generated was the same as that of light. In 1931 (during the centenary celebration of Maxwell's birth), Max Planck said:

> [Maxwell's equations remain] … for all time one of the greatest triumphs of human intellectual endeavor.

Albert Einstein said:

> [The work of Maxwell was] … the most profound and the most fruitful that physics has experienced since the time of Newton.

In a medium characterized by dielectric permittivity ε, the velocity of light waves is given by

$$v = \frac{1}{\sqrt{\varepsilon\,\mu}} = \frac{c}{n}, \tag{2.27}$$

with

$$n = \sqrt{\frac{\varepsilon\,\mu}{\varepsilon_0\mu_0}} \simeq \sqrt{\frac{\varepsilon}{\varepsilon_0}}, \tag{2.28}$$

which represents the refractive index of the medium, where we have assumed that $\mu \approx \mu_0$, which is true for all dielectrics.

The simplest wave (propagating, for instance, in the +*z* direction) is given by

$$\psi = A\, e^{i\omega\left(t - \frac{z}{v}\right)}, \tag{2.29}$$

which will always satisfy Eq. (2.22) for *any* value of the frequency ω. Equation (2.29) can also be written in the form

$$\psi = A\, e^{i(\omega t - kz)}, \tag{2.30}$$

where $k = \omega / v$ is known as the wave vector. Using Eq. (2.23), k can also be recast as

$$k = \omega\sqrt{\varepsilon\,\mu} = \frac{\omega}{c} n\,. \tag{2.31}$$

2.3 Plane Waves in a Homogenous Dielectric

A general 3D plane wave (for instance, propagating in the direction of **k**) is represented by the following equation:

$$\psi = A\, e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}. \tag{2.32}$$

If we substitute the preceding equation in Eq. (2.22), we find that

$$k_x^2 + k_y^2 + k_z^2 = k^2 = \frac{\omega^2}{v^2}. \tag{2.33}$$

For plane waves propagating in the direction of **k**, the electric and magnetic fields can be written in the form

$$\mathbf{E} = \mathbf{E}_0 \, e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}, \tag{2.34}$$

and

$$\mathbf{H} = \mathbf{H}_0 \, e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}, \tag{2.35}$$

where $\mathbf{E}_0$ and $\mathbf{H}_0$ are space- and time-independent vectors, but may, in general, be complex. The equation $\nabla \cdot \mathbf{E} = 0$ readily gives us

$$-i\left(k_x E_{0x} + k_y E_{0y} + k_z E_{0z}\right) = 0, \tag{2.36}$$

implying that

$$\mathbf{k} \cdot \mathbf{E} = 0. \tag{2.37}$$

Similarly, the equation $\nabla \cdot \mathbf{H} = 0$ gives us

$$\mathbf{k} \cdot \mathbf{H} = 0. \tag{2.38}$$

The preceding two equations tell us that **E** and **H** are at right angles to **k**. Now,

$$(\nabla \times \mathbf{E})_x = \left(\frac{\partial E_z}{\partial y} - \frac{\partial E_y}{\partial z}\right) = -i\left(k_y E_{0z} - k_z E_{0y}\right) e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}$$

$$= -i\,(\mathbf{k} \times \mathbf{E})_x. \tag{2.39}$$

Thus, Eq. (2.11) gives us

$$-i\,(\mathbf{k} \times \mathbf{E})_x = -i\,\omega\mu\, H_x \quad \Rightarrow \quad H_x = \frac{(\mathbf{k} \times \mathbf{E})_x}{\omega\mu}. \tag{2.40}$$

Similarly, we can obtain expressions for H_y and H_z and for the vector equation

$$\mathbf{H} = \frac{\mathbf{k} \times \mathbf{E}}{\omega\mu}. \tag{2.41}$$

And similarly, Eq. (2.12) gives us

$$\mathbf{E} = -\frac{\mathbf{k} \times \mathbf{H}}{\omega \varepsilon}, \tag{2.42}$$

showing that **H, E**, and **k** are at right angles to each other such that the vector **E** × **H** is parallel to **k**. As we will show later, the direction of energy propagation is along the Poynting vector:

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}, \tag{2.43}$$

which will be in the direction of **k**. If we substitute for **H** from Eq. (2.41) in Eq. (2.42), we obtain

$$\begin{aligned} \mathbf{E} &= -\frac{1}{\omega^2 \varepsilon \mu} \; [\mathbf{k} \times (\mathbf{k} \times \mathbf{E})] \\ &= -\frac{1}{\omega^2 \varepsilon \mu} \; [\mathbf{k}(\mathbf{k} \cdot \mathbf{E}) - \mathbf{E} \cdot (\mathbf{k} \cdot \mathbf{k})], \end{aligned} \tag{2.44}$$

where we have used the vector identity

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B} (\mathbf{A} \cdot \mathbf{C}) - \mathbf{C} (\mathbf{A} \cdot \mathbf{B}). \tag{2.45}$$

Since $\mathbf{k} \cdot \mathbf{E} = 0$, Eq. (2.44) reduces to

$$\mathbf{E} = \frac{k^2}{\omega^2 \varepsilon \mu} \; \mathbf{E}. \tag{2.46}$$

Thus,

$$k = \omega \sqrt{\varepsilon \mu}, \tag{2.47}$$

which is the same as Eq. (2.31). Without any loss of generality, if we assume the propagation to be along the *z* axis, then,

$$\mathbf{k} = \hat{\mathbf{z}} \, k, \tag{2.48}$$

where $\hat{\mathbf{z}}$ represents the unit vector along the *z* direction. Thus,

$$\mathbf{k} \cdot \mathbf{r} = k z, \tag{2.49}$$

and

$$e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})} = e^{i(\omega t - kz)}. \tag{2.50}$$

We assume the electric vector to be along the x axis, so the magnetic vector will be along the y axis because $\mathbf{E} \times \mathbf{H}$ should be in the direction of propagation z. Thus, we may write

$$\mathbf{E} = \hat{\mathbf{x}}\, E_0\, e^{i(\omega t - kz)}, \tag{2.51}$$

$$\mathbf{H} = \hat{\mathbf{y}}\, H_0\, e^{i(\omega t - kz)}, \tag{2.52}$$

with

$$H_0 = \frac{k}{\omega\mu} E_0. \tag{2.53}$$

The actual electric and magnetic fields are the real part of the exponentials appearing on the right-hand side of Eqs. (2.51) and (2.52):

$$\mathbf{E} = \hat{\mathbf{x}}\, E_0\, \cos[(\omega t - kz)], \tag{2.54}$$

$$\mathbf{H} = \hat{\mathbf{y}}\, H_0 \cos[(\omega t - kz)], \tag{2.55}$$

where we have assumed E_0 and H_0 to be real. The plane wave as represented by Eq. (2.51) [or by Eq. (2.54)] is said to be linearly polarized (or x polarized) because the electric vector is always along the x axis and the magnetic vector is always along the y axis [Figs. 2.1 and 2.2(a)]. Similarly, for a y-polarized wave, the electric vector is always along the y axis, as shown in Fig. 2.2(b).

From the preceding equations, we can draw the following inferences for plane waves:

(i) Both **E** and **H** are at right angles to the direction of propagation. Thus, the waves are transverse (see Fig 2.1).

(ii) The vectors **E** and **H** are at right angles to each other such that the direction of the vector $\mathbf{E} \times \mathbf{H}$ is along the direction of propagation **k**. Thus, if the direction of propagation is along the z axis and if **E** is assumed to point in the x direction, then **H** will point in the y direction (see Fig. 2.1).

(iii) Since $k/\omega\mu$ is a real number, the electric and magnetic vectors are in phase. Thus, if at any instant **E** is zero, then **H** is also zero; similarly, when **E** attains its maximum value, **H** also attains its maximum value, and so on.

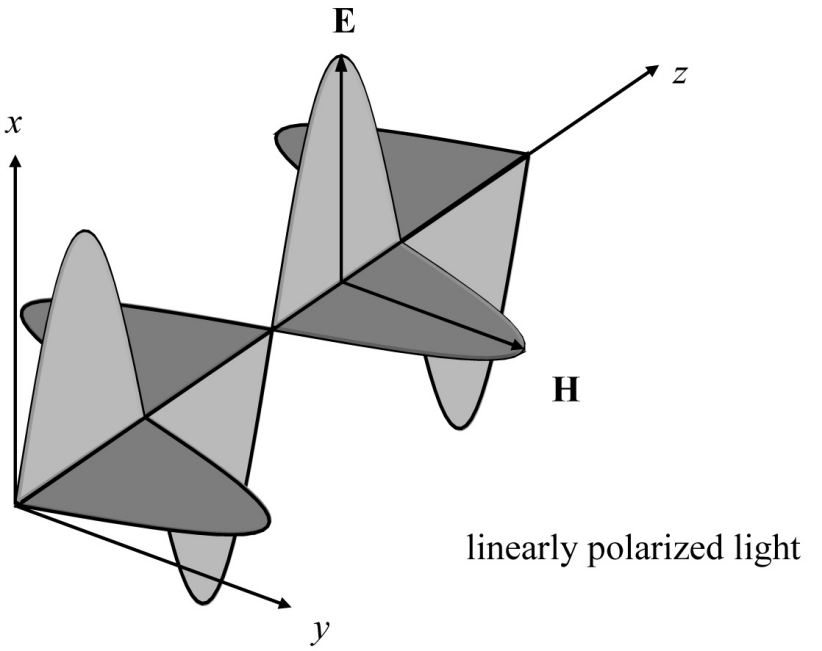


Figure 2.1 An *x*-polarized electromagnetic wave propagating in the *z* direction.

(iv) The electric and magnetic waves are interdependent—neither can exist without the other. Physically, an electric field varying in time produces a magnetic field varying in space and time; this changing magnetic field produces an electric field varying in space and time, and so on. This mutual generation of electric and magnetic fields results in the propagation of the electromagnetic wave.

(v) Since Maxwell's equations are linear in **E** and **H**, if ($\mathbf{E}_1$, $\mathbf{H}_1$) and ($\mathbf{E}_2$, $\mathbf{H}_2$) are two independent solutions of Maxwell's equations, then ($\mathbf{E}_1 + \mathbf{E}_2$, $\mathbf{H}_1 + \mathbf{H}_2$) will also be a solution of Maxwell's equations. This is the superposition principle according to which the resultant displacement produced by two independent disturbances is the vector sum of the displacements independently produced by the disturbances. Thus, the superposition principle is a consequence of the linearity of Maxwell's equations. If, for example, the field associated with the electromagnetic wave is so high that the dielectric permittivity ε depends on **E** itself, then Maxwell's equations will become nonlinear, and the superposition principle will not remain valid. Indeed, when we discuss any nonlinear phenomenon, the superposition principle does not hold.

(vi) The plane wave as represented by Eq. (2.51) is said to be linearly polarized because the electric vector is always along the *x* axis; similarly, the magnetic vector is always along the *y* axis (see Fig 2.1). In the next chapter, we will discuss different states of polarization by superposing linearly polarized waves.

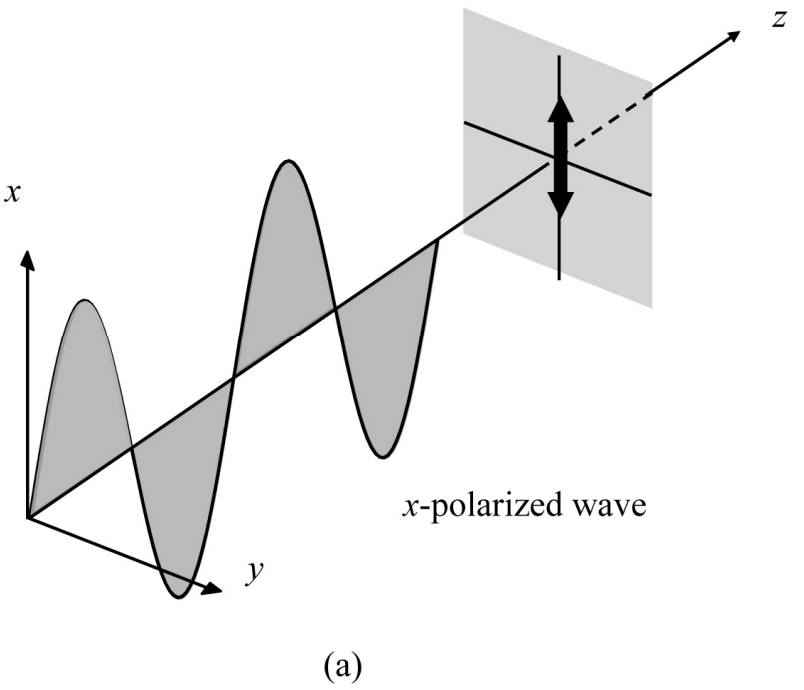


Figure 2.2(a) The electric field variation of an *x*-polarized wave.

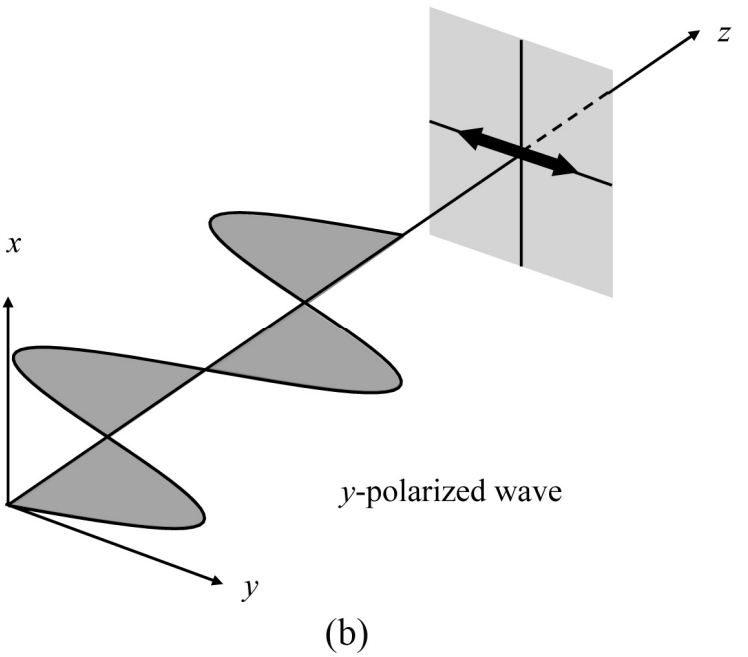


Figure 2.2(b) The electric field variation of a *y*-polarized wave.

2.4 The Poynting Vector

Let us consider Eqs. (2.3) and (2.4):

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.56}$$

and

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}. \tag{2.57}$$

Now,

$$\nabla \cdot (\mathbf{E} \times \mathbf{H}) = \mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{H})$$

$$= -\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} - \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} - \mathbf{J} \cdot \mathbf{E}. \tag{2.58}$$

For a linear material,

$$\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} = \mu \mathbf{H} \cdot \frac{\partial \mathbf{H}}{\partial t} + \varepsilon \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t}$$

$$= \frac{1}{2}\mu \frac{\partial}{\partial t}\left(\mathbf{H} \cdot \mathbf{H}\right) + \frac{1}{2}\varepsilon \frac{\partial}{\partial t}\left(\mathbf{E} \cdot \mathbf{E}\right)$$

$$= \frac{1}{2}\frac{\partial}{\partial t}\left(\mathbf{B} \cdot \mathbf{H} + \mathbf{D} \cdot \mathbf{E}\right). \tag{2.59}$$

Thus, Eq. (2.58) can be rewritten in the form

$$\text{div}\, \mathbf{S} + \frac{\partial u}{\partial t} = -\mathbf{J} \cdot \mathbf{E}, \tag{2.60}$$

where

$$\mathbf{S} \equiv \mathbf{E} \times \mathbf{H} \tag{2.61}$$

is known as the Poynting vector and

$$u = \frac{1}{2}\left(\mathbf{B} \cdot \mathbf{H} + \mathbf{D} \cdot \mathbf{E}\right). \tag{2.62}$$

Equation (2.59) is valid even for anisotropic media because, in the principal axis coordinate system,

$$\mathbf{E}\cdot\frac{\partial \mathbf{D}}{\partial t}=\frac{1}{2}\varepsilon_x\frac{\partial E_x^2}{\partial t}+\frac{1}{2}\varepsilon_y\frac{\partial E_y^2}{\partial t}+\frac{1}{2}\varepsilon_z\frac{\partial E_z^2}{\partial t}$$
$$=\frac{1}{2}\frac{\partial}{\partial t}\left[D_xE_x+D_yE_y+D_zE_z\right]=\frac{1}{2}\frac{\partial}{\partial t}(\mathbf{D}\cdot\mathbf{E}). \tag{2.63}$$

Equation (2.60) resembles the equation of continuity. For a physical interpretation of various terms in Eq. (2.60), we note that if a charge q (moving with velocity v) is acted on by an electromagnetic field, then the work done by the field in moving it through a distance $d\mathbf{s}$ would be $\mathbf{F}\ d\mathbf{s}$, where $\mathbf{F}$ is the force on the charge due to the electric and magnetic fields, given by

$$\mathbf{F}=q(\mathbf{E}+\mathbf{v}\times\mathbf{B})\,. \tag{2.64}$$

Thus, the work done per unit time would be

$$\mathbf{F}\cdot\frac{d\mathbf{s}}{dt}=\mathbf{F}\cdot\mathbf{v}$$
$$=[q\,\mathbf{E}+q\,\mathbf{v}\times\mathbf{B}]\cdot\mathbf{v}$$
$$=q\,\mathbf{E}\cdot\mathbf{v}. \tag{2.65}$$

If there are N charged particles per unit volume, each carrying a charge q, then the work done per unit volume would be

$$Nq\,\mathbf{v}\cdot\mathbf{E}=\mathbf{J}\cdot\mathbf{E}\,, \tag{2.66}$$

where $\mathbf{J}$ represents the current density. The energy appears in the form of kinetic (or heat) energy of the charged particles. Thus, the term $\mathbf{J}\cdot\mathbf{E}$ represents the familiar Joule loss; therefore, the quantity $\mathbf{J}\cdot\mathbf{E}$ appearing on the right-hand side of Eq. (2.66) represents the rate at which energy is produced per unit volume per unit time. We compare Eq. (2.60) with the equation of continuity, which is written in the form

$$\text{div}\ \mathbf{J}\ +\ \frac{\partial\rho}{\partial t}=0\,, \tag{2.67}$$

where ρ represents the charge density and $\mathbf{J}$ the current density; i.e., $\mathbf{J}\ d\mathbf{a}$ represents the amount of charge crossing the area $d\mathbf{a}$ per unit time. We may now interpret Eq. (2.60) as an equation of continuity for energy, with u representing the energy per unit volume, and the quantities ½ $\mathbf{D}\cdot\mathbf{E}$ and ½ $\mathbf{B}\cdot\mathbf{H}$ representing the electrical and magnetic energies per unit volume, respectively. Further, we

may interpret $\mathbf{S} \cdot d\mathbf{a}$ as the electromagnetic energy crossing the area $d\mathbf{a}$ per unit time. For plane waves in a dielectric, we may write

$$\mathbf{E} = \hat{\mathbf{x}}\ E_0 \cos(\omega t - kz)$$

and

$$\mathbf{H} = \hat{\mathbf{y}}\ H_0 \cos(\omega t - kz) = \hat{\mathbf{y}}\ \frac{k}{\omega\mu} E_0 \cos(\omega t - kz) \quad . \tag{2.68}$$

Thus,

$$\begin{aligned} \mathbf{S} &= \mathbf{E} \times \mathbf{H} \\ &= \hat{\mathbf{z}} \frac{k}{\omega\mu}\ E_0^2 \cos^2(\omega t - kz)\,, \end{aligned} \tag{2.69}$$

which implies that the energy flow is in the z direction (which represents the direction of propagation of the wave) and that energy in the amount of $\left(k/\omega\mu\right) E_0^2 \cos^2(\omega t - kz)$ crosses a unit area (perpendicular to the z axis) per unit time. For optical beams, $\omega \approx 10^{15}\ \mathrm{s}^{-1}$, the $\cos^2$ term fluctuates with extreme rapidity, and any detector would record only an average value. Since

$$\begin{aligned} \left\langle \cos^2(\omega t - kz) \right\rangle &= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{+T} \cos^2(\omega t - kz)\, dt \\ &= \frac{1}{2}, \end{aligned} \tag{2.70}$$

we obtain

$$\left\langle \mathbf{S} \right\rangle = \hat{\mathbf{z}}\ \frac{k}{2\omega\mu} E_0^2\,, \tag{2.71}$$

where $\langle \ldots \rangle$ denotes the time average of the quantity inside the angular brackets. If we now use

$$\mu = \mu_0 = \frac{1}{c^2 \varepsilon_0}\,, \tag{2.72}$$

and

$$\frac{k}{\omega} = \frac{1}{v} = \frac{n}{c}, \tag{2.73}$$

we obtain

$$\langle \mathbf{S} \rangle = \hat{\mathbf{z}}\ \frac{1}{2}\ \varepsilon_0 c\ n\ E_0^2. \tag{2.74}$$

Further, if we use the constitutive relations and perform time averaging, we obtain the following expressions for the energy density:

$$u = \frac{1}{4}\ \varepsilon E_0^2 + \frac{1}{4}\ \mu H_0^2 \tag{2.75}$$

or

$$u = \frac{1}{2}\ \varepsilon E_0^2, \tag{2.76}$$

where we used Eq. (2.53) to substitute H_0 in terms of E_0. The intensity of the beam will therefore be given by

$$I = v\left(\frac{1}{2}\ \varepsilon E_0^2\right) = \frac{c}{n}\left(\frac{1}{2}\ \varepsilon_0 n^2 E_0^2\right) \tag{2.77}$$

or

$$I = \frac{1}{2}\ \varepsilon_0 c\ n\ E_0^2, \tag{2.78}$$

which is consistent with Eq. (2.74).

Example 2.1 Obtain the intensity and the energy density for a 5-mW laser beam with 1-mm radius propagating through air.

Solution: The intensity of the laser beam is given by

$$I = \frac{5\times10^{-3}}{\pi\left(10^{-3}\right)^2} \approx 1.6\times10^3 \ \text{W/m}^2.$$

Thus,

$$E_0 = \sqrt{\frac{2I}{\varepsilon_0 c\, n}} = \sqrt{\frac{2\times 1.6\times 10^3}{8.854\times 10^{-12}\times 3\times 10^8\times 1}} \approx 1100 \text{ V/m},$$

and the energy density will be

$$u = \frac{1}{2}\,\varepsilon E_0^2 = \frac{1}{2}\times 8.854\times 10^{-12}\times (1100)^2 \approx 5.3\times 10^{-6} \text{ J/m}^3.$$

Note that $I = c\,u$.

Example 2.2 Obtain the intensity and the energy density at a distance of 2 m for a 200-W source emitting light uniformly in all directions.

Solution: The intensity at a distance of 2 m is given by

$$I = \frac{200}{4\pi(2)^2} \approx 4 \text{ W/m}^2.$$

Thus,

$$E_0 = \sqrt{\frac{2I}{\varepsilon_0 c\, n}} = \sqrt{\frac{2\times 4}{8.854\times 10^{-12}\times 3\times 10^8\times 1}} \approx 55 \text{ V/m},$$

and the energy density is

$$u = \frac{1}{2}\,\varepsilon E_0^2 = \frac{1}{2}\times 8.854\times 10^{-12}\times (55)^2 \approx 1.3\times 10^{-6} \text{ J/m}^3.$$

Bibliography

Born, M. and E. Wolf, *Principles of Optics*, 7th ed., Cambridge University Press, New York (1999).

Ghatak, A., *Optics*, McGraw Hill, New York (2009).

Ghatak, A. and K. Thyagarajan, *Introduction to Fiber Optics*, Cambridge University Press, New York (1998).

Griffiths, D. J., *Introduction to Electrodynamics*, 2nd ed., Prentice-Hall of India, New Delhi (1989).

Jackson, J. D., *Classical Electrodynamics*, 3rd ed., J. Wiley & Sons, New York (1998).

Chapter 3
Basic Concepts of Polarization

3.1 Introduction

As discussed in Chapter 2, the electric and magnetic field vectors of a plane electromagnetic (EM) wave lie in a plane transverse to the direction of propagation. If the electric field varies randomly with time, then the wave is said to be *unpolarized*, while if it varies in a predictable manner, the wave is said to be *polarized*. The nature of variation of the electric field with time defines the state of polarization (SOP) of the EM wave. For example, the variation may be such that the tip of the electric field vector may move on a line, on a circle, or, in general, on an ellipse. Accordingly, the SOP of the wave is said to be linear, circular, or elliptical. In the following discussion, we provide the basic characteristics of polarized waves.

3.2 Various Polarization States

3.2.1 Linear polarization

The electric field vector of a linearly polarized wave is the easiest of the three to describe mathematically. For example, if the line over which the tip of the electric field vector moves is parallel to the x axis, then the electric field can be described as

$$\mathbf{E} = \hat{\mathbf{x}}\ a\cos(\omega t - k z + \delta_0)\,, \tag{3.1}$$

where $k = (2\pi / \lambda_0)n$, with n being the refractive index of the medium, λ_0 the free-space wavelength, z the direction of propagation, a and δ_0 the amplitude and the initial phase, respectively, and $\hat{\mathbf{x}}$ the unit vector along the x axis. Such a wave is also known as *plane polarized* because the electric field of such a wave always lies in a longitudinal plane (the x-z plane in this case). Without any loss of generality, we may assume that $a > 0$ and $-\pi \le \delta_0 \le \pi$. Thus, we may write

$$\mathbf{E} = -\hat{\mathbf{x}}\ a\cos\left(\omega t - k z - \frac{\pi}{3}\right) = \hat{\mathbf{x}}\ a\cos\left(\omega t - k z + \frac{2\pi}{3}\right) \tag{3.2}$$

and

$$\mathbf{E} = -\hat{\mathbf{x}}\, a\cos\left(\omega t - k z + \frac{\pi}{6}\right) = \hat{\mathbf{x}}\, a\cos\left(\omega t - k z - \frac{5\pi}{6}\right), \qquad (3.3)$$

and so on. In the preceding equations, a is always positive.

In the more general case of a linearly polarized wave, the electric field vector may make an angle θ with the x axis. Such a wave can be described as the superposition of two orthogonal and linearly polarized components along the x and y axes, with appropriate amplitudes and initial phases, given by

$$E_x = a\cos\theta\, \cos(\omega t - k z + \delta_0), \qquad (3.4)$$

$$E_y = a\sin\theta\, \cos(\omega t - k z + \delta_0). \qquad (3.5)$$

The preceding two components will have same initial phase if $0 < \theta < \pi/2$ and will differ in phase by π if $\pi / 2 < \theta < \pi$.

For example, for the linearly polarized light shown in Fig. 3.1(a), $\theta = \pi/6$, and the two components are given by

$$E_x = \frac{a\sqrt{3}}{2}\cos(\omega t - k z + \delta_0), \qquad (3.6)$$

and

$$E_y = \frac{a}{2}\cos(\omega t - k z + \delta_0), \qquad (3.7)$$

which have same initial phase. On the other hand, for the linearly polarized light shown in Fig. 3.1(b), $\theta = 5\pi/6$, and the two components are

$$E_x = -\frac{a\sqrt{3}}{2}\cos(\omega t - k z + \delta_0) = \frac{a\sqrt{3}}{2}\cos(\omega t - k z + \delta_0 - \pi) \qquad (3.8)$$

and

$$E_y = \frac{a}{2}\cos(\omega t - k z + \delta_0), \qquad (3.9)$$

which differ in phase by π.

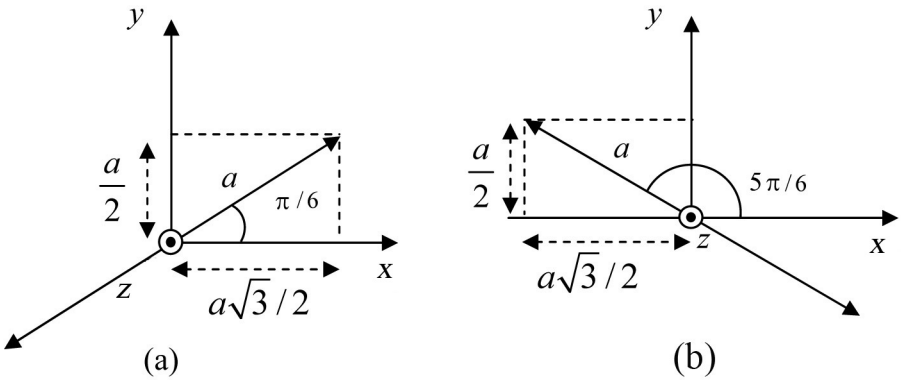


Figure 3.1 A linearly polarized wave making an angle (a) $\pi/6$ and (b) $5\pi/6$ with the *x* axis.

3.2.2 Circular polarization

The SOP of a given plane wave is known to be *circular* if at a given point in space, the tip of the electric field vector moves on a circle with increase in time. Such a wave can be expressed as a superposition of two orthogonal and linearly polarized components (for instance, along the *x* and *y* axes) with equal amplitudes but having a phase difference of $\pi/2$, as given here:

$$E_x = a\cos\left(\omega t - kz + \delta_0\right), \tag{3.10}$$

$$\begin{aligned} E_y &= a\cos\left(\omega t - kz + \delta_0 \mp \frac{\pi}{2}\right) \\ &= \pm\, a\sin\left(\omega t - kz + \delta_0\right). \end{aligned} \tag{3.11}$$

It can easily be shown that the amplitude of the resulting electric field

$$\mathbf{E} = \hat{\mathbf{x}}\ a\cos\left(\omega t - kz + \delta_0\right) \pm \hat{\mathbf{y}}\ a\sin\left(\omega t - kz + \delta_0\right) \tag{3.12}$$

is $a\left(=\sqrt{E_x^2 + E_y^2}\right)$, which does not change with time. However, the direction of the electric field keeps changing such that the tip of the **E** vector rotates on a circle of radius *a*, described by

$$E_x{}^2 + E_y{}^2 = a^2. \tag{3.13}$$

It should be noted that the positive (+) and negative (−) signs in Eq. (3.12) correspond to different states of polarization, as discussed in the following sections.

3.2.2.1 Left-circular polarization (LCP)

If the *y* component lags behind the *x* component in phase, then the resultant electric field vector rotates counterclockwise with increase in time. The corresponding SOP is said to be *left circular* and is given by

$$E_x = a\cos\left(\omega t - kz + \delta_0\right), \tag{3.14}$$

$$E_y = a\cos\left(\omega t - kz + \delta_0 - \frac{\pi}{2}\right) = a\sin\left(\omega t - kz + \delta_0\right). \tag{3.15}$$

3.2.2.2 Right-circular polarization (RCP)

If the *y* component leads the *x* component, then the electric field vector rotates clockwise with time. Such an SOP is said to be *right circular* and is given by

$$E_x = E_0\cos\left(\omega t - kz + \delta_0\right), \tag{3.16}$$

$$E_y = a\cos\left(\omega t - kz + \delta_0 + \frac{\pi}{2}\right) = -a\sin\left(\omega t - kz + \delta_0\right). \tag{3.17}$$

The rotation is taken clockwise and counterclockwise with respect to an observer who is looking into the source, i.e., against the direction of propagation. The initial phase δ_0 in the case of a circularly polarized wave represents the angle made by the electric field vector with the *x* axis at $z = 0$ and $t = 0$, which is $+\delta_0$ for the left rotation and $-\delta_0$ for the right rotation (Fig. 3.2).

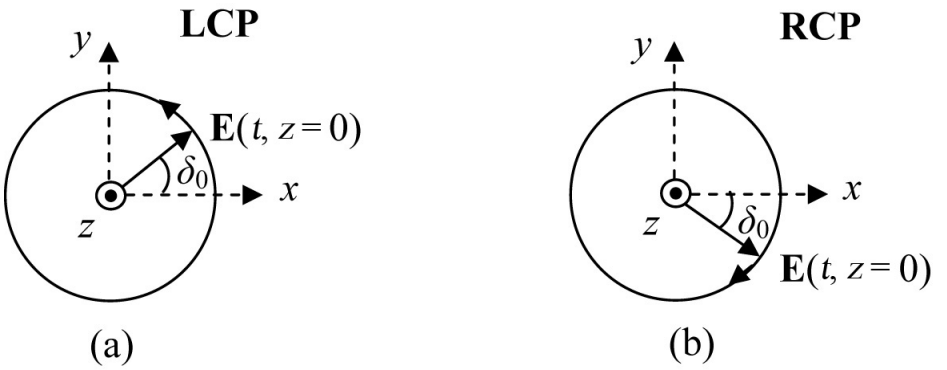


Figure 3.2 Schematic of (a) left- and (b) right-circularly polarized waves with initial phase δ_0. The propagation is along the +*z* direction.

3.2.3 Elliptical Polarization

The SOP of a given plane wave is known to be *elliptical* if at a given point in space (for instance, $z = 0$), the tip of the electric field vector moves on an ellipse, known as the *polarization ellipse*. The description of such an SOP is similar to that of a circular SOP if described by the two linear components polarized along the axes of the ellipse (for example, ξ and η). We will assume the major axis to be along the ξ axis. Further, if the lengths of the major and minor axes of the polarization ellipse are $2a$ and $2b$, respectively, such an SOP can be described in terms of linear components E_ξ and E_η, as follows:

$$E_\xi = a\cos\left(\omega t + \delta_0\right), \tag{3.18a}$$

$$E_\eta = \pm b\sin\left(\omega t + \delta_0\right), \tag{3.18b}$$

where the upper (+) and lower (−) signs correspond to the left and right sense of rotation (assuming a and b to be positive), and δ_0 is again the initial phase. It may be mentioned here that in this case also, δ_0 determines the initial orientation of the electric field vector, however, in the manner shown in Fig. 3.3. The dashed circle in Fig. 3.3 is of radius a so that $E_\xi(t = 0) = \mathrm{OA} = \mathrm{OQ}\cos\delta_0 = a\cos\delta_0$.

Further, since point P lies on the ellipse (solid curve),

$$\left(\frac{E_\xi}{a}\right)^2 + \left(\frac{E_\eta}{b}\right)^2 = 1. \tag{3.19}$$

Substituting $E_\xi = a\cos\delta_0$ in Eq. (3.19), we obtain

$$E_\eta(t = 0) = PA = \pm b\,\sin\delta_0. \tag{3.20}$$

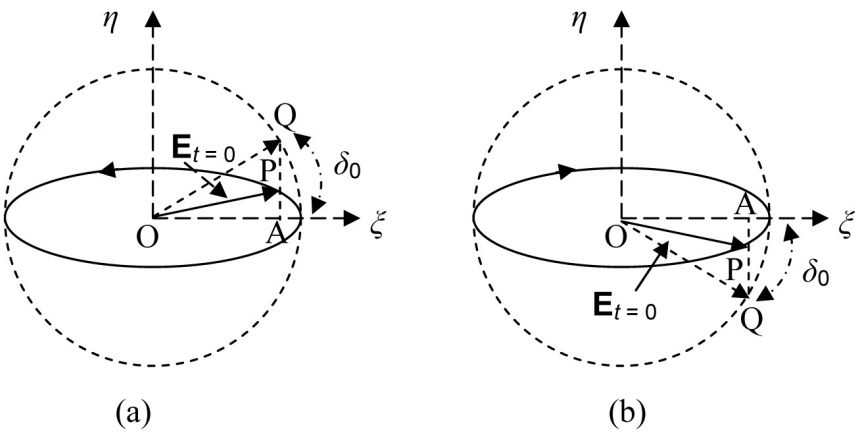


Figure 3.3 The direction of the electric field for (a) left and (b) right elliptical states of polarization, described by Eqs. (3.18a) and (3.18b).

Knowing E_ξ and E_η, we can obtain any two orthogonal linearly polarized components describing the given elliptical SOP. For example, if the ξ axis is at an angle θ with respect to the x axis (Fig. 3.4), the x and y components of the given SOP can be obtained simply by rotating the ξ-η coordinates clockwise by an angle θ, and are given by

$$E_x = E_\xi \cos\theta - E_\eta \sin\theta, \tag{3.21}$$

$$E_y = E_\xi \sin\theta + E_\eta \cos\theta. \tag{3.22}$$

In this way, we can obtain any pair of orthogonal linear components describing a given elliptical SOP.

Conversely, when two orthogonal and linearly polarized components of arbitrary amplitudes and phase difference superimpose, the resultant SOP is, in general, elliptical, as discussed in the following section. We will always assume that $0 \le \theta \le \pi$.

3.3 Superposition of Two Orthogonal Linear Polarizations

Let the x and y components of the electric field at $z = 0$ be given by

$$E_x = a_x \cos\left(\omega t + \delta_x\right), \tag{3.23}$$

$$E_y = a_y \cos\left(\omega t + \delta_y\right). \tag{3.24}$$

In the preceding equations, we are assuming that both a_x and a_y are positive. In order to obtain the equation of the polarization ellipse, we eliminate the (ωt) dependence from the preceding equations and write Eq. (3.24) as

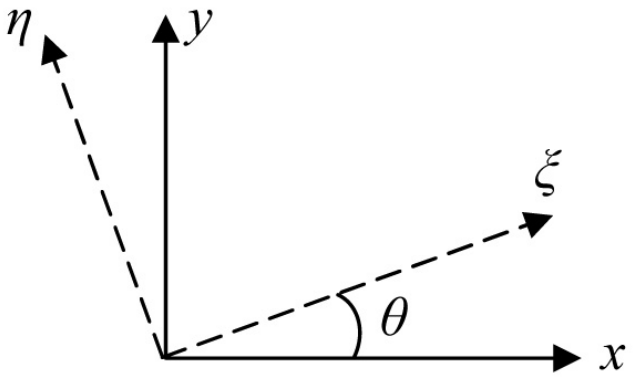


Figure 3.4 Relative directions of the x-y and ξ-η axes.

$$\frac{E_y}{a_y} = \cos\left(\omega t + \delta_x + \delta\right)$$

$$= \cos\left(\omega t + \delta_x\right)\cos\delta - \sin\left(\omega t + \delta_x\right)\sin\delta, \quad (3.25)$$

where $\delta = (\delta_y - \delta_x)$ represents the phase difference between the x and y components. Using Eq. (3.23) in the preceding equation, it can easily be shown that E_x, and E_y satisfy the following equation:

$$\frac{E_y}{a_y} = \frac{E_x}{a_x}\cos\delta - \left[1 - \left(\frac{E_x}{a_x}\right)^2\right]^{\frac{1}{2}}\sin\delta, \quad (3.26)$$

which can be put in the form

$$\left(\frac{E_x}{a_x}\right)^2 + \left(\frac{E_y}{a_y}\right)^2 - \frac{2E_x}{a_x}\frac{E_y}{a_y}\cos\delta = \sin^2\delta. \quad (3.27)$$

It can be shown that Eq. (3.27) represents an ellipse whose orientation and ellipticity can be obtained as follows: Let the lengths of the major and minor axes of the polarization ellipse be $2a$ and $2b$, respectively, and (ξ, η) be the axes parallel to the major and minor axes of the ellipse. As discussed earlier [Eqs. (3.18a) and (3.18b)], the given SOP can then be represented in terms of linear components along the ξ and η axes as follows:

$$E_\xi = a\cos\left(\omega t + \delta_0\right), \quad (3.28)$$

$$E_\eta = \pm b\sin\left(\omega t + \delta_0\right). \quad (3.29)$$

It should be remembered that the positive/negative sign in the preceding equation represents left/right rotation of the electric field vector, respectively. Assuming that the ξ axis makes an angle θ with the x axis, the field components E_ξ and E_η can also be obtained from E_x and E_y by rotating the coordinate system (ξ, η) counterclockwise by an angle θ as

$$E_\xi = E_x\cos\theta + E_y\sin\theta, \quad (3.30)$$

and

$$E_\eta = -E_x\sin\theta + E_y\cos\theta. \quad (3.31)$$

Substituting E_x and E_y from Eqs. (3.23) and (3.24) and E_ξ and E_η from Eqs. (3.28) and (3.29) in the preceding equations, we obtain

$$a\cos\left(\omega t+\delta_0\right)=a_x\cos\left(\omega t+\delta_x\right)\cos\theta+a_y\cos\left(\omega t+\delta_y\right)\sin\theta, \quad (3.32)$$

$$\pm b\sin\left(\omega t+\delta_0\right)=-a_x\cos\left(\omega t+\delta_x\right)\sin\theta+a_y\cos\left(\omega t+\delta_y\right)\cos\theta. \quad (3.33)$$

Comparing the coefficients of cos (ωt) and sin (ωt) in Eq. (3.32), we obtain

$$a\cos\delta_0=a_x\cos\delta_x\cos\theta+a_y\cos\delta_y\sin\theta, \quad (3.34)$$

$$a\sin\delta_0=a_x\sin\delta_x\cos\theta+a_y\sin\delta_y\sin\theta. \quad (3.35)$$

Similarly, by comparing the coefficients of sin (ωt) and cos (ωt) in Eq. (3.33), we obtain

$$\pm b\cos\delta_0=a_x\sin\delta_x\sin\theta-a_y\sin\delta_y\cos\theta, \quad (3.36)$$

$$\pm b\sin\delta_0=-a_x\cos\delta_x\sin\theta+a_y\cos\delta_y\cos\theta. \quad (3.37)$$

By adding the squares of Eqs. (3.34) and (3.35) and then of (3.36) and (3.37), we obtain

$$a^2=a_x^2\cos^2\theta+a_y^2\sin^2\theta+a_xa_y\sin 2\theta\cos\delta \quad (3.38)$$

$$b^2=a_x^2\sin^2\theta+a_y^2\cos^2\theta-a_xa_y\sin 2\theta\cos\delta. \quad (3.39)$$

By adding the two preceding equations, we obtain

$$a^2+b^2=a_x^2+a_y^2, \quad (3.40)$$

which is expected, as both sides of the preceding equation represent the intensity of the wave. Further, by dividing Eq. (3.36) by Eq. (3.34), and Eq. (3.37) by Eq. (3.35), we obtain

$$\pm\frac{b}{a}=\frac{a_x\sin\delta_x\sin\theta-a_y\sin\delta_y\cos\theta}{a_x\cos\delta_x\cos\theta+a_y\cos\delta_y\sin\theta}=\frac{-a_x\cos\delta_x\sin\theta+a_y\cos\delta_y\cos\theta}{a_x\sin\delta_x\cos\theta+a_y\sin\delta_y\sin\theta}. \quad (3.41)$$

The preceding equation can be easily simplified to show that the orientation of the ellipse is given by

$$\tan 2\theta = \frac{2a_x a_y}{\left(a_x^2 - a_y^2\right)} \cos\delta \,. \tag{3.42}$$

Defining an auxiliary angle α such that

$$\alpha = \tan^{-1} \frac{a_y}{a_x}, \tag{3.43}$$

Eq. (3.42) can be put in the following form:

$$\tan 2\theta = \tan 2\alpha \cos\delta \,, \tag{3.44}$$

where $\alpha(0 \le \alpha \le \pi/2)$ represents the angle between the x axis and the diagonal of the rectangle (with axes parallel to the x and y axes) enveloping the ellipse [see Fig. 3.5(a)].

We now multiply Eq. (3.34) by Eq. (3.36), and Eq. (3.35) by Eq. (3.37), and add to give

$$\pm ab = -a_x a_y \sin\delta \,. \tag{3.45}$$

Using Eq. (3.40), the preceding equation can be put in the following form:

$$\mp \frac{2ab}{a^2 + b^2} = \frac{2a_x a_y}{\left(a_x^2 + a_y^2\right)} \sin\delta \,. \tag{3.46}$$

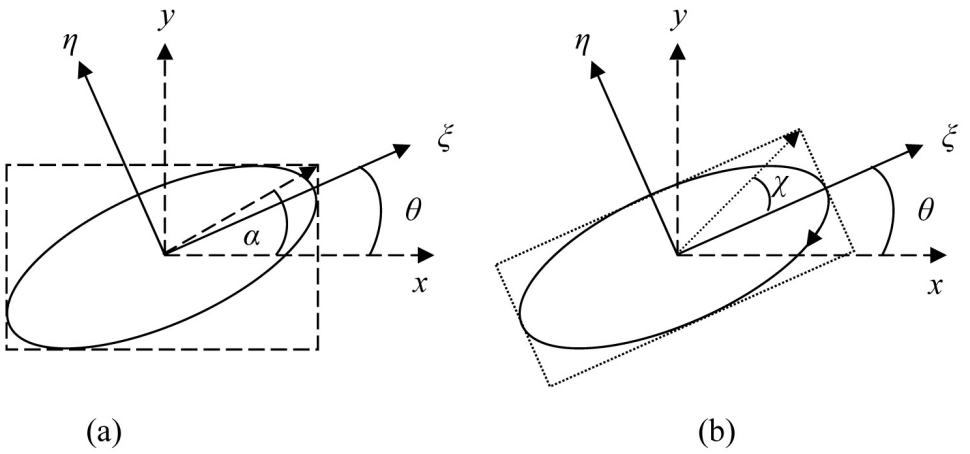


Figure 3.5 Geometrical representations of angles (a) α, θ, and (b) χ.

Defining another auxiliary angle χ such that

$$\tan \chi = \mp \frac{b}{a}, \tag{3.47}$$

Eq. (3.46) can be put in the form

$$\sin 2\chi = \sin 2\alpha \sin \delta . \tag{3.48}$$

The angle χ ($-\pi/4 \le \chi \le \pi/4$), known as the *ellipticity angle*, is a measure of both ellipticity as well as the sense of rotation of the resultant SOP. A positive (negative) value of χ corresponds to a negative (positive) sign in Eq. (3.29) and therefore represents a right (left) sense of rotation. Furthermore, $\chi = 0$ means that $b/a = 0$; therefore, the polarization state represented by Eqs. (3.28) and (3.29) is a linearly polarized wave. Geometrically, χ is the angle between the major axis of the ellipse and the diagonal of the rectangle having sides parallel to major and minor axes and enveloping the ellipse [see Fig. 3.5(b)].

3.3.1 Method for obtaining the polarization state

The characteristics of the SOP described by Eqs. (3.23) and (3.24) are obtained as follows:

First we calculate $\delta = (\delta_y - \delta_x)$ and

$$\alpha = \tan^{-1} \frac{a_y}{a_x}, 0 \le \alpha \le \frac{\pi}{2}.$$

Next,

(i) orientation θ is given by

$$\tan 2\theta = \tan 2\alpha \cos\delta. \tag{3.49}$$

(ii) Ellipticity (b/a) and the sense of rotation are given by angle χ, which is determined by

$$\sin 2\chi = \sin 2\alpha \sin \delta , \quad -\frac{\pi}{4} \le \chi \le \frac{\pi}{4}. \tag{3.50}$$

Here, a positive (negative) value of χ represents a right (left) sense of rotation. It should be noted that if both a_x and a_y are positive (as assumed), sin 2α will always be positive; therefore, for sin $\delta > 0$, χ will be positive (i.e., right rotation), while for sin $\delta < 0$, χ will be negative (i.e., left rotation). Some specific cases are discussed as follows:

(i) $\delta = \pm n\pi$, with n = 0, 1, 2, …. In this case, sin δ = 0. This means that χ = 0, and the resultant SOP is linear. Further,

for n = 0, 2, 4, …, cos δ = 1; therefore,

$$\tan 2\theta = \tan 2\alpha \Rightarrow \theta = \alpha . \tag{3.51}$$

For n = 1, 3, 5, …, $\cos\delta = -1$; therefore,

$$\tan 2\theta = -\tan 2\alpha \Rightarrow \theta = -\alpha . \tag{3.52}$$

(ii) $\delta = (\pm n + ½)\ \pi$, with n = 0, 1, 2, …. In this case, cos δ = 0, which means that $\tan 2\theta = 0 \Rightarrow 2\theta = 0$ or π, i.e., θ = 0 or π/2. It can easily be shown that

$$\theta = 0 \text{ for } a_y < a_x \text{, i.e., } \alpha < \pi/4,$$

$$= \pi/2 \text{ for } a_y > a_x \text{, i.e., } \alpha > \pi/4 .$$

For $a_x = a_y$, the resultant SOP is circular for which θ is not defined. Furthermore, sin $\delta = (-1)^n$, which means that $\chi = (-1)^n\ \alpha$. Thus, the resultant SOP is right rotating for n = 0, 2, 4, …, and left rotating for n = 1, 3, 5, … .

(iii) $a_x = a_y = a$. In this case, $\alpha = \tan^{-1}\dfrac{a_y}{a_x} = \dfrac{\pi}{4}$ and $\tan 2\alpha = \infty$.

Substituting this in $\tan 2\theta = \tan 2a \cos\delta$ gives

$$\theta = \frac{\pi}{4} \quad \text{for } \cos\delta > 0 \quad \left\{0 < \delta < \frac{\pi}{2}\right\}$$

$$= \frac{3\pi}{4} \quad \text{for } \cos\delta < 0 \quad \left\{\frac{3\pi}{2} < \delta < 2\pi\right\}.$$

For general values of a_x, a_y, and δ, the resultant polarization states are shown schematically in Fig. 3.6.

It should be noted that for a given value of α and δ, there are two possible ways of selecting the value of 2θ that satisfy Eq. (3.49). For example, if the right-hand side of Eq. (3.49) is positive, we can select 2θ either in the first quadrant (0 to π/2) or in the third quadrant (π to 3 π/2). Similarly, if the right-hand side of Eq. (3.49) is negative, we can select 2θ in either the second quadrant (π/2 to π) or in the fourth quadrant (3π/2 to 2π).

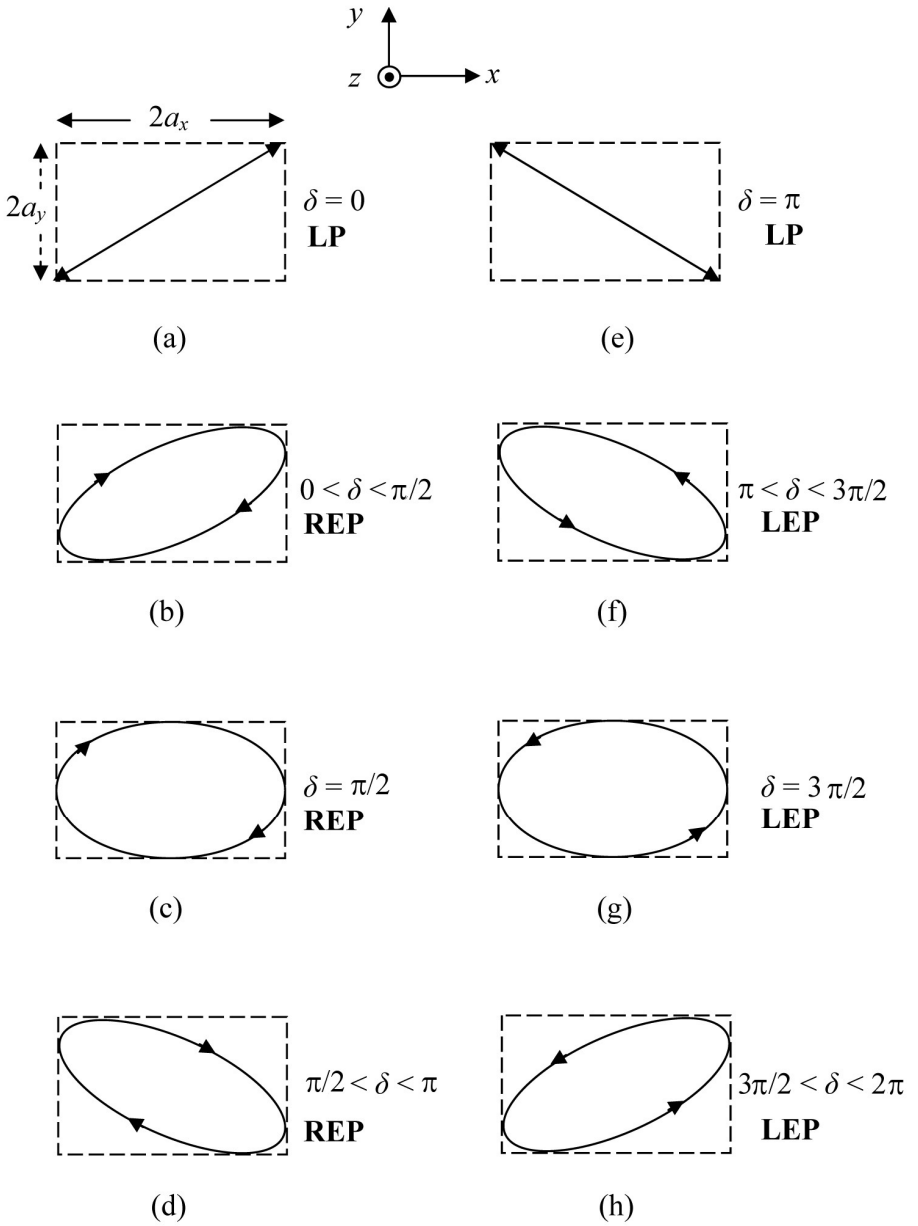


Figure 3.6 Polarization states described by Eqs. (3.23) and (3.24) for various values of δ. The propagation is along the +*z* direction.

The procedure to select the correct value of 2θ is given in the following:

(i) For $0 \le \alpha < \dfrac{\pi}{4}$ (i.e., $a_y < a_x$), $0 < 2\theta < \dfrac{\pi}{2}$, if $\tan 2\theta > 0$,

and

$$3\pi/2 < 2\theta < 2\pi, \text{ if } \tan 2\theta < 0.$$

(ii) For $\pi/4 < \alpha < \pi/2$ (i.e., $a_y > a_x$), $\pi/2 < 2\theta < \pi$, if $\tan 2\theta < 0$,

and

$$\pi < 2\theta < 3\pi/2, \text{ if } \tan 2\theta > 0.$$

(iii) For $\alpha + \pi/4$, $\tan 2\theta = \infty$, i.e., $2\theta = \pi/2$ or $3\pi/2$.

Thus,

$$2\theta = \pi/2, \text{ if } \cos\delta > 0, \text{ and } 2\theta = 3\pi/2, \text{ if } \cos\delta < 0.$$

Figure 3.7 shows the correct regions in which 2θ should be selected for given values of α and $\cos\delta$. Now we provide a few simple examples that use the preceding method to determine the SOP for which the x and y components are given.

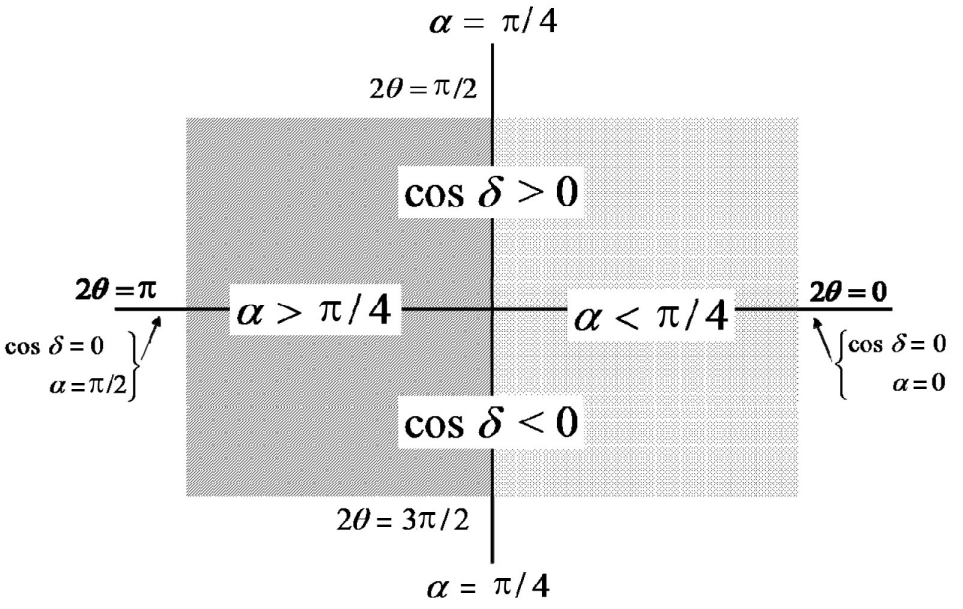


Figure 3.7 Correct range of 2θ for a given value of α and δ.

Example 3.1 Obtain the polarization state as described by

$$E_x = \frac{\sqrt{3}}{2}\cos\left(\omega t - kz\right),$$
$$E_y = \frac{1}{2}\cos\left(\omega t - kz + \pi/4\right).$$

Solution: First we calculate α and δ:

$$\alpha = \tan^{-1}\frac{a_y}{a_x} = \tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{6} \text{ and } \delta = \delta_y - \delta_x = \pi/4.$$

Substituting α and δ in $\tan 2\theta = \tan 2\alpha \cos\delta$, we obtain $\tan 2\theta = 1.224745$. In this case, α is less than $\pi/4$ and $\tan 2\theta$ is positive, so we select 2θ in the first quadrant, i.e., between π and $\pi/2$, which is 50.77 deg. This gives the orientation of the major axis as $\theta = 25.39$ deg.

Next, we substitute α and δ in $\sin 2\chi = \sin 2\alpha \sin\delta$, which gives $\chi = 18.88$ deg. This means that the given SOP represents a right- (as χ is positive) elliptically polarized wave whose major axis is oriented at $\theta = 25.39$ deg, with $b/a = \tan(\chi) = 0.34$. Thus, the polarization state will be somewhat similar to that shown in Fig. 3.6(b).

Example 3.2 Obtain the polarization state as described by

$$E_x = \frac{\sqrt{3}}{2}\cos\left(\omega t - kz\right),$$
$$E_y = \frac{1}{2}\cos\left(\omega t - kz + 3\pi/4\right).$$

Solution: First we calculate α and δ:

$$\alpha = \tan^{-1}\frac{a_y}{a_x} = \tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{6} \text{ and } \delta = \delta_y - \delta_x = 3\pi/4.$$

Substituting α and δ in $\tan 2\theta = \tan 2\alpha \cos\delta$, we obtain $\tan 2\theta = -1.224745$. From here, we can select 2θ in either the second quadrant or the fourth quadrant, which is 129.23 deg or 309.23 deg, respectively. The latter is correct in this case, because α is less than $\pi/4$ and $\tan 2\theta$ is negative, so we select 2θ in the fourth quadrant (Fig. 3.7), i.e., 309.23 deg. This gives the orientation of the major axis as $\theta = 154.62$ deg.

Next we substitute α and δ in $\sin 2\chi = \sin 2\alpha \sin\delta$, which gives $\chi = 18.88$ deg. This means that the given SOP represents a right- (as χ is positive)

elliptically polarized wave whose major axis is oriented at θ = 154.62 deg, with $b/a = \tan(\chi) = 0.34$. Thus, the polarization state will be somewhat similar to that shown in Fig. 3.6(d).

Example 3.3 Obtain the polarization state as described by

$$E_x = \frac{1}{2}\cos\left(\omega t - kz\right),$$
$$E_y = \frac{\sqrt{3}}{2}\cos\left(\omega t - kz + \pi/4\right).$$

Solution: First we calculate α and δ:

$$\alpha = \tan^{-1}\frac{a_y}{a_x} = \tan^{-1}\left(\sqrt{3}\right) = \frac{\pi}{3} \text{ and } \delta = \delta_y - \delta_x = \pi/4.$$

Substituting α and δ in $\tan 2\theta = \tan 2\alpha \cos\delta$, we obtain $\tan 2\theta = -1.224745$. From here, we can select 2θ in either the second or the fourth quadrant, which is 129.23 deg or 309.23 deg, respectively. Unlike the previous example, in this case, the former is correct, because α is greater than $\pi/4$ and $\tan 2\theta$ is negative, so we select 2θ in the second quadrant (Fig. 3.7). This gives the orientation of the major axis as θ = 64.62 deg.

Next we substitute α and δ in $\sin 2\chi = \sin 2\alpha \sin\delta$, which gives χ = 18.88 deg. This means that the given SOP represents a right- (as χ is positive) elliptically polarized wave whose major axis is oriented at θ = 64.62 deg with $b/a = \tan(\chi) = 0.34$. Thus, the polarization state will be somewhat similar to that shown in Fig. 3.6(b).

Example 3.4 Obtain the polarization state as described by

$$E_x = \frac{1}{2}\cos\left(\omega t - kz\right),$$
$$E_y = \frac{\sqrt{3}}{2}\cos\left(\omega t - kz + 3\pi/4\right).$$

Solution: First we calculate α and δ:

$$\alpha = \tan^{-1}\frac{a_y}{a_x} = \tan^{-1}\left(\sqrt{3}\right) = \frac{\pi}{3} \text{ and } \delta = \delta_y - \delta_x = 3\pi/4.$$

Substituting α and δ in $2\theta = \tan 2\alpha \cos\delta$, we obtain $\tan 2\theta = 1.224745$. From here, we can select 2θ in either the first or the third quadrant, which is 50.70 deg or 230.77 deg, respectively. Unlike Example 3.1, in this case, the latter is correct,

because α is greater than $\pi/4$ and $\tan 2\theta$ is positive, so we select 2θ in the third quadrant (Fig. 3.7). This gives the orientation of the major axis as $\theta = 115.38$ deg.

Next, we substitute α and δ in $2\chi = \sin 2\alpha \sin \delta$, which gives $\chi = 18.88$ deg. This means that the given SOP represents a right- (as χ is positive) elliptically polarized wave whose major axis is oriented at $\theta = 115.38$ deg, with $b/a = \tan(\chi) = 0.34$. Thus, the polarization state will be somewhat similar to that shown in Fig. 3.6(d).

Example 3.5 Obtain the polarization state as described by

$$E_x = a\cos\left(\omega t - kz\right),$$
$$E_y = a\cos\left(\omega t - kz + \pi/4\right).$$

Solution: First we calculate α and δ:

$$\alpha = \tan^{-1}\frac{a_y}{a_x} = \frac{\pi}{4} \text{ and } \delta = \delta_y - \delta_x = \pi/4.$$

Substituting α and δ in $\tan 2\theta = \tan 2\alpha \cos \delta$, we obtain $\tan 2\theta = \infty$. Since $\cos \delta$ is positive, we select 2θ as $\pi/2$; therefore, $\theta = \pi/4$. Now, substituting α and δ in $\sin 2\chi = \sin 2\alpha \sin \delta = \sin \delta$, we obtain $\chi = \pi/8$. This means that the given SOP represents a right- (as χ is positive) elliptically polarized wave whose major axis is oriented at $\theta = \pi/4$ with $b/a = \tan(\pi/8) = 0.41$. Thus, the polarization state will be somewhat similar to that shown in Fig. 3.6(b).

Example 3.6 Obtain the polarization state as described by

$$E_x = a\cos\left(\omega t - kz\right),$$
$$E_y = a\cos\left(\omega t - kz + 3\pi/4\right).$$

Solution: First we calculate α and δ:

$$\alpha = \tan^{-1}\frac{a_y}{a_x} = \frac{\pi}{4} \text{ and } \delta = \delta_y - \delta_x = 3\pi/4.$$

Substituting α and δ in $\tan 2\theta = \tan 2\alpha \cos \delta$, we obtain $\tan 2\theta = \infty$. Since $\cos\delta$ is negative, unlike the previous example, here we select 2θ as $3\pi/2$; therefore, $\theta = 3\pi/4$. Now, substituting α and δ in $\sin 2\chi = \sin 2\alpha \sin\delta = 1/\sqrt{2}$, we obtain $\chi = \pi/8$ (as χ lies between $-\pi/4$ and $\pi/4$). This means that the given SOP represents a right- (as χ is positive) elliptically polarized wave whose major axis is oriented at $\theta = 3\pi/4$ with $b/a = \tan(\pi/8) = 0.41$. Thus, the polarization state will be somewhat similar to that shown in Fig. 3.6(d).

3.4 Retarders or Wave Plates

The discussion presented in the preceding section shows that any SOP can be thought of as a superposition of two orthogonal linearly polarized components with appropriate amplitudes and phase difference; therefore, by changing the phase difference between these components, one can generate a new SOP. This can be achieved using a device known as a *retarder*.

A retarder is an optical component that splits the input SOP into two orthogonally polarized components and introduces a desired phase difference (*retardation*) between them using the phenomenon of *birefringence*. Inside a retarder, one of the orthogonal components travels with a low refractive index and is known as the *fast* component, while the other traveling with a higher refractive index is known as the *slow* component. The medium of the retarder is known as *linearly*, *circularly*, or *elliptically* birefringent if the two orthogonal components are linearly, circularly, or elliptically polarized. Accordingly, the retarder is also known to be a linear, circular, or elliptical retarder. Linear retarders introducing a phase difference of $\pi/2$ or π, known as a *quarter-wave plate* (QWP) or a *half-wave plate* (HWP), respectively, are commonly used. The construction and working principle of wave plates will be discussed in more detail in the next chapter. The following examples demonstrate how the SOP of a given beam is changed by a wave plate.

Example 3.7 Obtain the output SOP when a left-circularly polarized beam is passed through a QWP.

Solution: At the input end of the QWP ($z = 0$), the given SOP can be written in terms of its x and y components as

$$E_x(z=0) = a\cos(\omega t)$$
$$E_y(z=0) = a\sin(\omega t).$$

Let us assume that the x and y axes are parallel to the fast and slow axes of the QWP. In this case, inside the QWP, the x and y components will travel with refractive indices n_f and n_s, respectively ($n_f < n_s$), and the two components at the output end ($z = d$) of the QWP will be given by

$$E_x(z=d) = a\cos(\omega t - k_0 n_f\, d) \text{ and}$$
$$E_y(z=d) = a\sin(\omega t - k_0 n_s\, d)$$
$$= a\sin\{\omega t - k_0 n_f\, d - k_0(n_s - n_f)\, d\},$$

which shows that the y component will be slowed down in phase by $\delta = k_0(n_s - n_f)d$ as compared to the x component. For a QWP, $\delta = \pi/2$; therefore, the output SOP is described by

$$E_x = a\cos\left(\omega t - k_0 n_{\mathrm{f}}\, d\right) \text{ and}$$
$$E_y = a\sin\left(\omega t - k_0 n_{\mathrm{f}}\, d - \pi/2\right)$$
$$= -a\cos\left(\omega t - k_0 n_{\mathrm{f}}\, d\right),$$

which represents a linearly polarized wave making an angle of –45 deg with the x axis (Fig. 3.8). Similarly, it can be shown that a QWP will:

(i) Convert a right-circularly polarized beam to a linearly polarized wave making an angle of +45 deg with the fast axis.
(ii) Convert a linearly polarized beam making an angle of +45 deg (or –45 deg) with the fast axis to a left- (or right-) circularly polarized beam.

Example 3.8 Obtain the output SOP when a left-elliptically polarized beam is passed through a QWP whose fast axis is aligned with the major axis of the ellipse.

Solution: Let the incident SOP be represented by a polarization ellipse with semi-major and semi-minor axes as a and b and with its major axis at angle θ with the x axis. At the input end ($z = 0$) of the QWP, the SOP can be written in terms of its major and minor axes (for instance, ξ and η) components as

$$E_\xi(z=0) = a\cos\left(\omega t\right),$$
$$E_\eta(z=0) = b\sin\left(\omega t\right).$$

The ellipticity angle χ for the given beam is $\chi = -\varepsilon$, where $\varepsilon = \tan^{-1}(b/a)$. Inside the QWP, the ξ (fast) and η (slow) components will travel with refractive indices n_{f} and n_{s}, respectively ($n_f < n_s$), and the η component will be slowed down in phase by $\pi/2$ as compared to the ξ component. Therefore, the two components at the output end ($z = d$) of the QWP will be given by

$$E_\xi(z=d) = a\cos\left(\omega t - k_0 n_{\mathrm{f}}\, d\right) \text{ and}$$
$$E_\eta(z=d) = b\sin\left(\omega t - k_0 n_{\mathrm{f}}\, d - \pi/2\right)$$
$$= -b\cos\left(\omega t - k_0 n_{\mathrm{f}}\, d\right),$$

which represent a linearly polarized wave making an angle of $-\varepsilon = -\tan^{-1}(b/a)$ with the ξ axis (fast axis) and, therefore, an angle of $\theta - \varepsilon$ with the x axis (Fig. 3.9).

Similarly, it can be shown that if the incident SOP is right elliptical, the output SOP will be a linearly polarized wave, making an angle of $+\varepsilon$ with the fast axis and $\theta + \varepsilon$ with the x axis.

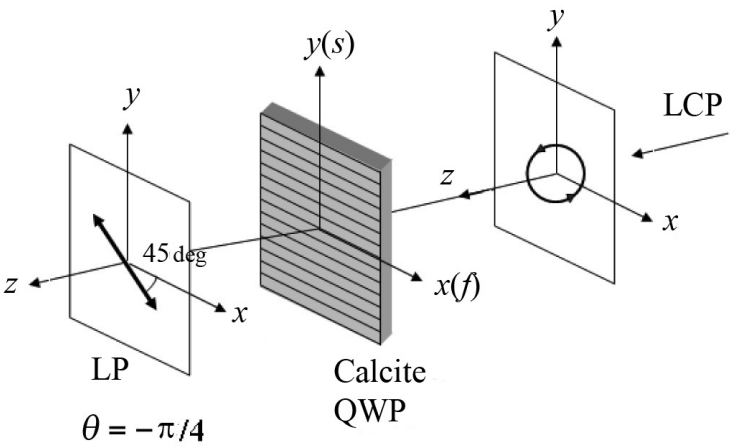


Figure 3.8 If a left-circularly polarized wave is incident on a QWP (whose fast axis is along the *x* axis), the output beam will be linearly polarized, making an angle θ (= $-\pi/4$) with the *x* axis.

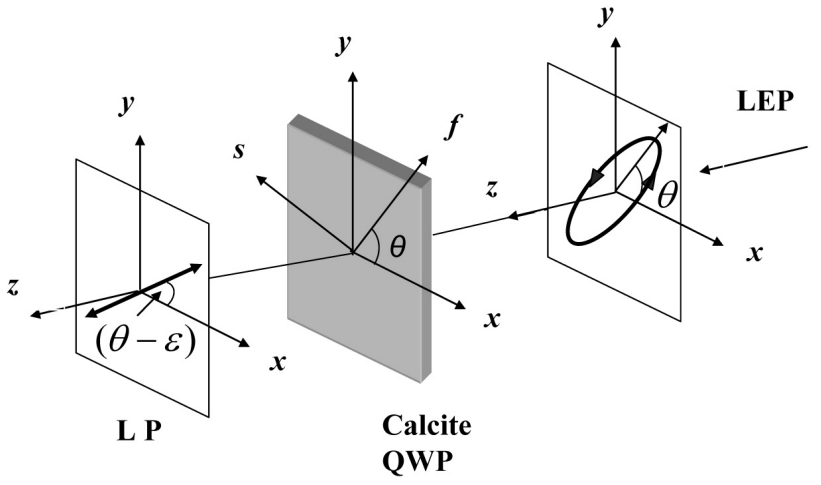


Figure 3.9 A left-elliptically polarized wave oriented at θ with ellipticity *b*/*a* will be converted by a QWP (whose fast axis is along θ) to a linearly polarized wave making an angle ($\theta - \varepsilon$) with the *x* axis; $\varepsilon = \tan^{-1}(b/a)$.

Example 3.9 Obtain the output SOP when a linearly polarized beam is passed through a QWP.

Solution: If the incident beam is polarized parallel to the fast (or slow) axis of the QWP, then the beam will pass through the QWP as a fast (or slow) component; therefore, its SOP will remain the same.

Let us now assume that the fast axis of the QWP makes an angle θ with the *x* axis and that the incident linearly polarized beam makes an angle ε with the fast axis. Thus, the incident light is polarized at an angle $\theta + \varepsilon$ with the *x* axis. If the

directions of the fast and slow axes are taken as ξ and η, the input SOP at the input end ($z = 0$) of the QWP can be represented as

$$E_\xi(z=0) = a\cos(\omega t)$$
$$E_\eta(z=0) = b\cos(\omega t),$$

where $a = A\cos\varepsilon$ and $b = A\sin\varepsilon$, with A representing the amplitude of the incident linearly polarized beam. Inside the wave plate, the η component will travel as a slow component; therefore, its phase will be decreased by $\pi/2$ as compared to the ξ component (see Example 3.7). Therefore, the two components at the output of the QWP can be written as

$$E_\xi(z=d) = a\cos(\omega t - k_0 n_f\, d) \text{ and}$$
$$E_\eta(z=d) = b\cos(\omega t - k_0 n_f\, d - \pi/2),$$
$$= b\sin(\omega t - k_0 n_f\, d),$$

respectively, which represent a left-elliptically polarized wave with ellipticity $b/a = \tan(\varepsilon)$, whose major axis is parallel to the fast axis of the QWP and, therefore, at an angle θ with the x axis (Fig. 3.10). Similarly, it can be shown that if the incident linearly polarized beam makes an angle $-\varepsilon$ with the fast axis of the QWP, the output SOP will be a right-elliptically polarized wave with ellipticity $b/a = \tan(\varepsilon)$, with the major axis again at an angle θ with the x axis.

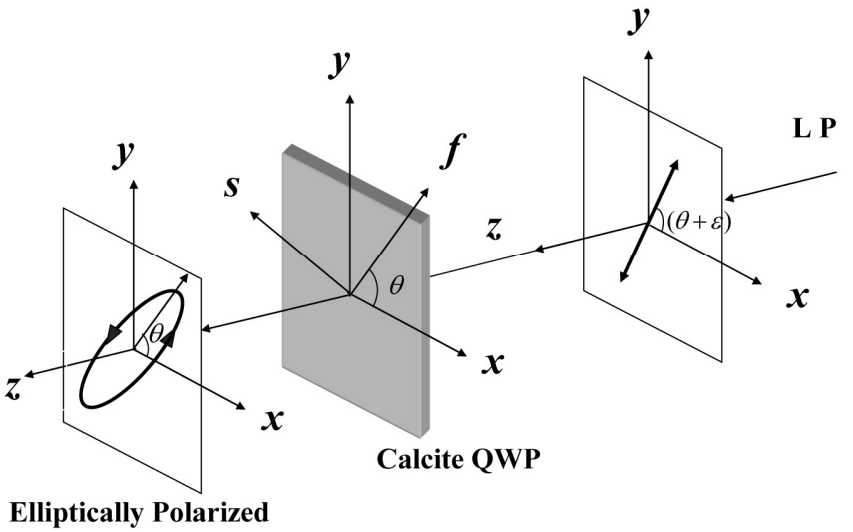


Figure 3.10 A linearly polarized wave at $(\theta + \varepsilon)$ will be converted by a QWP (whose fast axis is at θ) to a left-elliptically polarized wave oriented at θ with ellipticity $b/a = \tan(\varepsilon)$.

In the preceding example we assumed that $\varepsilon < \pi/4$. Choosing $\varepsilon = \pi/4$, it is evident from the preceding example that the output SOP will be left- (right-) circularly polarized if the incident SOP makes an angle of $+\pi/4$ ($-\pi/4$) with the fast axis of the QWP.

3.5 Polarizers

Polarizers are used to produce a desired polarization state. An ideal polarizer is an optical element that splits the incident polarization state into two orthogonally polarized components and allows only one component to pass through while stopping the other. The polarizer is said to be *linear*, *circular*, or *elliptical* if the passed component is linearly, circularly, or elliptically polarized.

It is clear from Example 3.9 that the most general SOP (elliptical) represented by an ellipse with ellipticity b/a and whose major axis is at an angle θ with the x axis can be produced by passing a linearly polarized beam through a QWP whose fast axis is at an angle θ with the x axis and whose input beam is polarized at an angle of $\pm \tan^{-1}(b/a)$ with the fast axis of the QWP. Thus, producing a linear SOP is the first step in producing an arbitrary SOP. In the following sections, we discuss methods for producing a linearly polarized wave.

3.5.1 Producing linearly polarized light

There are several ways to produce a linearly polarized wave; the following sections present some of these methods.

3.5.1.1 Using a Polaroid

A Polaroid is a thin, plastic-like sheet consisting of a long chain of molecules aligned in a particular direction. If an unpolarized light wave is incident on such a Polaroid, the associated electric field vectors along the direction of the aligned molecules are absorbed. Thus, the light wave will become linearly polarized with the electric vector oscillating along a direction perpendicular to the aligned molecules; this direction is known as the *pass axis* of the Polaroid.

3.5.1.2 Using reflection

Let us consider the incidence of an unpolarized plane wave on a plane surface separating two dielectrics of refractive indices n_1 and n_2. If the angle of incidence θ is such that

$$\theta = \theta_{\mathrm{p}} = \tan^{-1}\left(\frac{n_2}{n_1}\right), \tag{3.53}$$

the reflected beam will be linearly polarized, and its electric vector will be perpendicular to the plane of incidence (see Fig. 3.11).

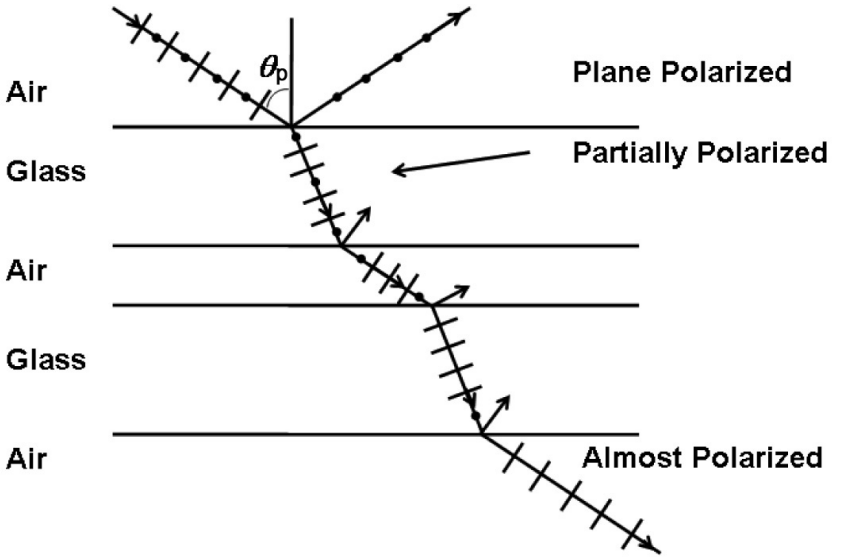


Figure 3.11 If an unpolarized beam is incident with an angle of incidence equal to θ_p, the reflected beam is plane polarized; its electric vector is perpendicular to the plane of incidence and the transmitted beam is partially polarized.

Equation (3.53) is referred to as *Brewster's law*, and θ_p is known as the *Brewster angle* or *polarizing angle*. At this angle of incidence, the reflected and transmitted rays are at right angles to each other. For the air–glass interface, $n_1 = 1$ and $n_2 = 1.5$, giving $\theta_p \approx 56.3$ deg. The transmitted beam is partially polarized. Therefore, if one uses a large number of reflecting surfaces, the transmitted light will also be linearly polarized.

3.5.1.3 Using double refraction

When an unpolarized beam enters an anisotropic crystal such as calcite, it splits into two components, one of which is linearly polarized, and the other orthogonally polarized; this phenomenon is known as *double refraction*, which will be discussed in detail in the next chapter.

If, by some method, we could eliminate one of the components, we would obtain a linearly polarized beam. A simple method for eliminating one of the beams uses the property of selective absorption; this property is also known as *dichroism*. A crystal such as tourmaline has a different coefficient of absorption for the two linearly polarized beams into which the incident beam splits. Consequently, one of the beams is absorbed quickly and the other component passes through without much attenuation. Thus, if an unpolarized beam is passed through a tourmaline crystal, the emergent beam will be linearly polarized. Another method of eliminating one of the polarized components is through total internal reflection, which will be discussed in the next chapter.

3.5.1.4 Linear polarization and Malus' law

Let us consider a linearly polarized beam propagating in the z direction, with its electric field at an angle θ with the x axis, incident on a linear polarizer P whose pass axis is parallel to the x axis (Fig. 3.12). If the amplitude of the incident electric field is E_0, the wave emerging from the polarizer P will be of amplitude $E_0 \cos\theta$ and will be polarized along the x axis. Therefore, the intensity of the emerging beam will be given by

$$I = I_0 \cos^2 \theta, \tag{3.54}$$

where I_0 represents the intensity of the incident beam. Equation (3.54) is referred to as *Malus' law*.

If the incident beam is not linearly polarized, we first express the incident SOP in terms of linear components parallel and perpendicular to the pass axis of polarizer P; the parallel component will pass through, while the perpendicular component will be stopped.

3.5.2 Producing circularly polarized light

It is evident from Example 3.9 that a linear polarizer followed by a QWP whose fast axis is at an angle of $-\pi/4$ ($\pi/4$) with the pass axis of the polarizer will produce a left- (right-) circularly polarized wave. However, it should be noted that the preceding combination of a linear polarizer and a QWP does not form an ideal left- (right-) circular polarizer, as it will not completely pass an incident left- (right-) circularly polarized beam, due to a 50% reduction in intensity by the linear polarizer. Instead, a QWP with its fast axis along the x axis, for instance, followed by a linear polarizer with its pass axis at an angle of $-\pi/4$ ($\pi/4$) with the x axis, followed by another QWP with its slow axis along the x axis, makes an ideal left- (right-) circular polarizer (Fig. 3.13). We will discuss this in more detail in Chapters 5 and 7.

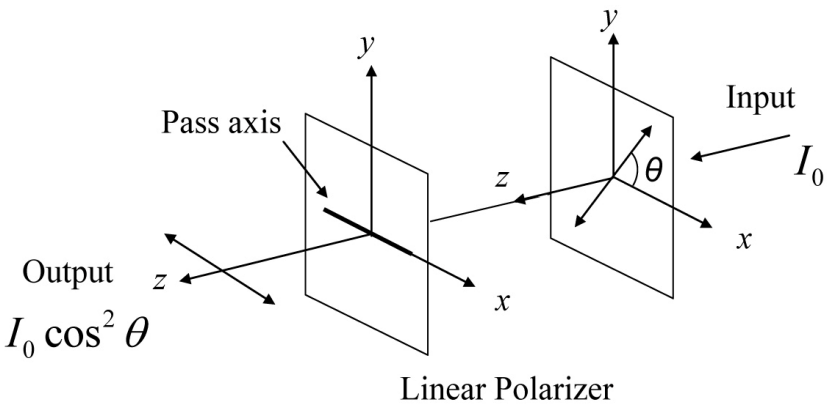


Figure 3.12 The output intensity will be $I = I_0 \cos^2 \theta$ when a linearly polarized wave of intensity I_0 oriented at θ is passed through a linear polarizer (whose pass axis is along the x axis).

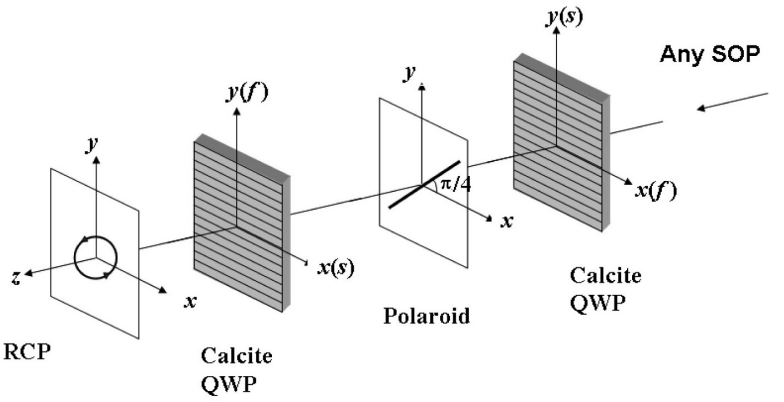


Figure 3.13 The combination of a QWP with its fast axis along the *x* axis, followed by a linear polarizer with its pass axis at an angle $\pi/4$ with the *x* axis, followed by another QWP with its slow axis along the *x* axis forms an ideal right-circular polarizer.

3.5.3 Producing elliptically polarized light

It is also clear from Example 3.9 that a linear polarizer followed by a QWP whose fast axis is at an angle θ with the *x* axis, while the pass axis of the polarizer is at an angle of ε ($-\varepsilon$) with respect to the fast axis of the QWP, will produce a left- (right-) elliptically polarized beam oriented at an angle θ with ellipticity $b/a = \tan \varepsilon$. Again, it should be noted that the preceding combination of a linear polarizer and a QWP does not form an ideal left- (right-) elliptical polarizer, as it will not completely pass the corresponding left- (right-) elliptically polarized beam, due to a 50% reduction in intensity by the polarizer. Instead, a QWP with its fast axis at θ, followed by a linear polarizer with its pass axis at an angle $-\varepsilon$ (ε) with respect to the fast axis of the QWP, then followed by another QWP with its slow axis at θ makes an ideal left- (right-) elliptical polarizer that produces a left- (right-) elliptically polarized beam oriented at an angle θ with ellipticity $b/a = \tan \varepsilon$ (Fig. 3.14). This point will be further discussed in Chapter 7.

3.6 Analysis of Polarized Light

Analysis of a given light wave with respect to its polarization state is an extremely important aspect of polarization optics. In this section, we discuss how to determine the SOP of a given light wave. A given light beam may be polarized, unpolarized, or partially polarized (a mixture of polarized and unpolarized). Thus, the various possible states of polarization for a given light beam are:

- linearly polarized (LP),
- circularly polarized (CP),
- elliptically polarized (EP),
- unpolarized (UP),
- combination of unpolarized and linearly polarized (UP + LP),
- combination of unpolarized and circularly polarized (UP + CP), and
- combination of unpolarized and elliptically polarized (UP + EP).

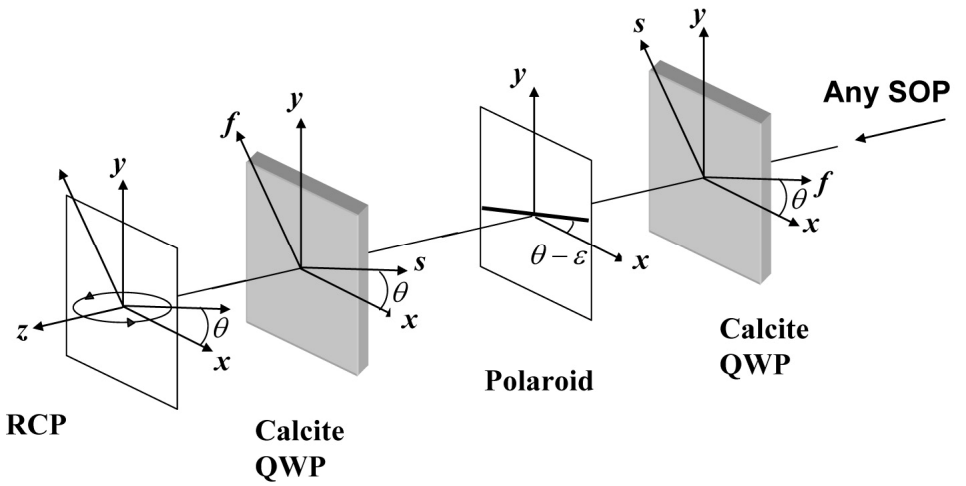


Figure 3.14 A combination of a QWP with its fast axis along θ, followed by a linear polarizer with its pass axis along $(\theta - \varepsilon)$, followed by another QWP with its slow axis along θ forms an ideal left-elliptical polarizer producing a left-elliptical SOP of orientation θ and ellipticity tan ε.

In order to determine the SOP of the given beam, we first pass it through a rotating linear analyzer (LA), e.g., a Polaroid, and record the output intensity variation as a function of the angle θ between the pass axis of the LA and the x axis. There are three possible outputs, as follows:

(i) The output varies periodically with complete extinction of light at two values of θ [Fig. 3.15(a)]. This clearly means that the input SOP is linearly polarized and oriented at $\theta = \theta_0$, corresponding to the maximum output intensity.

(ii) The output does not vary at all [Fig. 3.15(b)]. In this case, the possible states of polarization are either UP or CP, or a combination of UP and CP. In order to further distinguish them, we now insert a QWP between the source and the rotating analyzer and again record the output intensity variation as a function of θ. After passing through a QWP, the UP light remains unpolarized, while a left- (or right) CP wave is converted to a LP wave, making an angle of $-\pi/4$ (or $+\pi/4$) with the fast axis of the QWP (see Example 3.7). Thus, after inserting the QWP,

- if there is no variation of intensity, the input SOP will be unpolarized;
- if the intensity varies periodically with complete extinction of light at two values of θ, the SOP will be circularly polarized; and
- if the intensity varies periodically without complete extinction of light for any θ, the SOP will be a combination of UP and CP.

In the last two cases, the maximum intensity will occur at an angle of $-\pi/4$ (or $+\pi/4$) with respect to the fast axis of the QWP, indicating the left (or right) sense of rotation of the CP component.

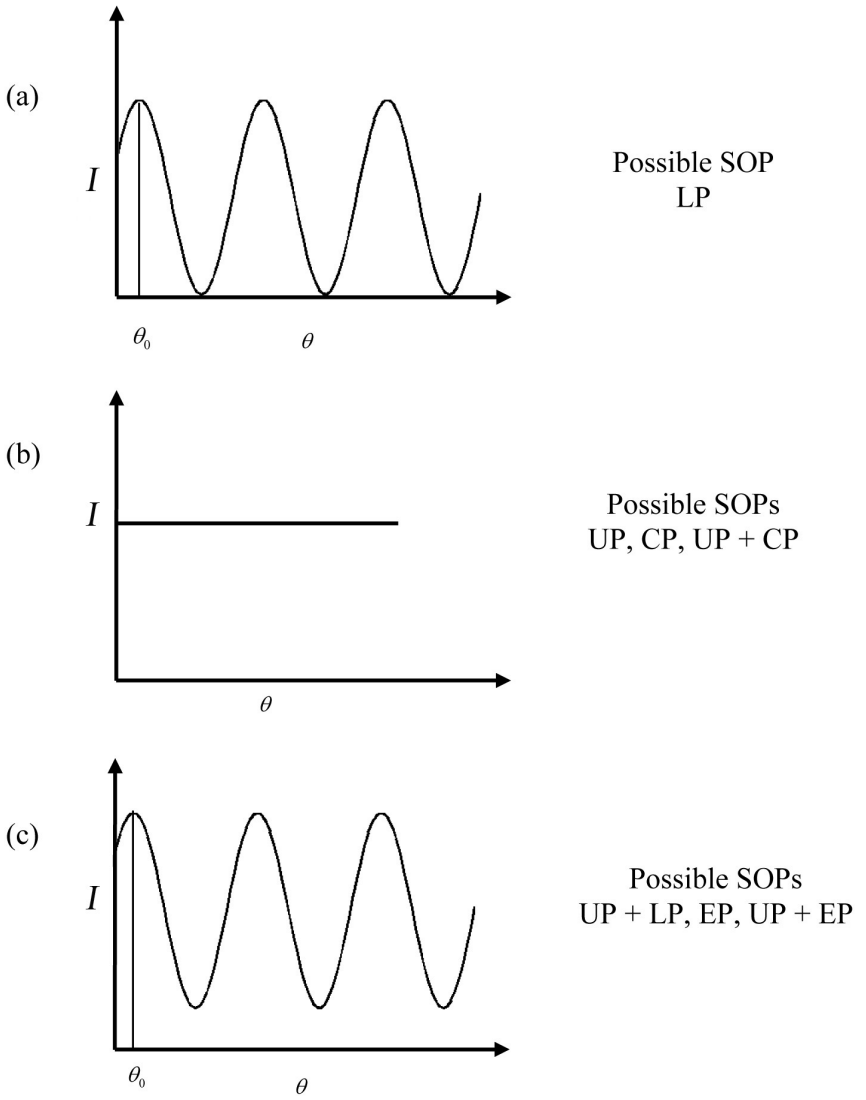


Figure 3.15 Various possible states of polarization and the corresponding variations of their output intensity when an incident light beam is passed through a rotating linear analyzer. (a) The output varies periodically with complete extinction of light at two values of θ. (b) The output does not vary at all. (c) The output varies periodically without complete extinction of light for any value of θ.

(iii) The output varies periodically without complete extinction of light for any value of θ [Fig. 3.15(c)]. In this case, the possible states of polarization are either (a) a combination of UP and LP, (b) EP, or (c) a combination of UP and EP. In order to further distinguish between these three possibilities, we first identify the orientation θ_0 corresponding to the maximum output intensity and then insert a QWP between the source and the rotating analyzer with its fast axis along θ_0. It is

evident that θ_0 will be the direction of polarization of the LP component for case (a) but will be the direction of the major axis of the polarization ellipse for cases (b) and (c). Thus, the QWP will not make any change in the incident SOP in case (a); however, in cases (b) and (c), a left- (or right-) EP wave will be converted to an LP wave, making an angle of $-\tan^{-1}(b/a)$ [or $+\tan^{-1}(b/a)$] with respect to the fast axis of the QWP (see Example 3.9). Thus, after inserting the QWP,

- if there is no change in the intensity variation, the input SOP will be a combination of UP and LP, with the linear component polarized along θ_0;
- if the intensity varies periodically with complete extinction of light at two values of θ, the SOP will be EP; and
- if the intensity varies periodically without complete extinction of light for any θ, the SOP will be a combination of UP and EP.

In the last two cases, the major axis of the EP component in the input SOP will be at θ_0, and its sense of rotation and ellipticity (b/a) can be obtained by measuring the new position of the maximum intensity, which will be $\theta_0 - \tan^{-1}(b/a)$ [or $\theta_0 + \tan^{-1}(b/a)$] if the EP component is left- (or right-) rotating, as discussed earlier.

Bibliography

Azzam, R. M. A. and N. M. Bashara, *Ellipsometry and Polarized Light*, North Holland Publishing Company, Amsterdam (1977).

Born, M. and E. Wolf, *Principles of Optics*, 7th ed., Cambridge University Press, New York (1999).

Collett, E., *Field Guide to Polarization*, SPIE Press, Bellingham, WA (2005). [doi:10.1117/3.626141].

Ghatak, A., *Optics*, McGraw-Hill, New York (2009).

Ghatak, A. and K. Thyagarajan, *Introduction to Fiber Optics*, Cambridge University Press, New York (1998).

Goldstein, D., *Polarized Light*, 2nd ed., Marcel Dekker, New York (2003).

Shurcliff, W. A., *Polarized Light: Production and Use*, Harvard University Press, Cambridge, MA (1962).

Theocaris, P. S, and E. E. Gdoutos, *Matrix Theory of Photoelasticity*, Springer Verlag, Berlin–Heidelberg (1979).

Chapter 4
Double Refraction and Applications

4.1 Introduction

When an unpolarized light beam enters an anisotropic medium such as calcite, it usually splits into two linearly polarized beams. One of the beams obeys Snell's law of refraction and is known as the *ordinary ray* (o ray). The second beam, which, in general, does not obey Snell's law, is known as the *extraordinary ray* (e ray). These two waves are linearly polarized and, in general, propagate with different wave and ray velocities (wave velocity is also referred to as phase velocity), and hence are characterized by different refractive indices. This phenomenon is known as *double refraction*, or *birefringence*, which is of great importance and is used to realize several polarization-based devices. In this chapter, we will discuss plane wave propagation in an anisotropic medium and some applications.

4.2 Anisotropic Media

The difference between an isotropic and an anisotropic medium lies in the relationship between the displacement vector **D** and the electric vector **E**. In an isotropic medium, **D** is in the same direction as **E**, and we can write

$$\mathbf{D} = \varepsilon \mathbf{E}, \tag{4.1}$$

where ε is the dielectric permittivity of the medium. On the other hand, in an anisotropic medium, **D** is not, in general, in the direction of **E**, and the relationship between **D** and **E** can be written in the form

$$\begin{aligned} D_x &= \varepsilon_{xx} E_x + \varepsilon_{xy} E_y + \varepsilon_{xz} E_z, \\ D_y &= \varepsilon_{yx} E_x + \varepsilon_{yy} E_y + \varepsilon_{yz} E_z, \\ D_z &= \varepsilon_{zx} E_x + \varepsilon_{zy} E_y + \varepsilon_{zz} E_z, \end{aligned} \tag{4.2}$$

where ε_{xx}, ε_{xy},… are constants. It can be shown that[1]

$$\varepsilon_{xy} = \varepsilon_{yx},\ \varepsilon_{xz} = \varepsilon_{zx}, \text{ and } \varepsilon_{yz} = \varepsilon_{zy}. \tag{4.3}$$

Further, we can always choose a coordinate system (i.e., the directions of the *x*, *y*, and *z* axes) inside the crystal such that

$$D_x = \varepsilon_x E_x,\ D_y = \varepsilon_y E_y, \text{ and } D_z = \varepsilon_z E_z. \tag{4.4}$$

This coordinate system is known as the *principal axis system*. The preceding equations show that if **E** is along one of the principal axes, **D** is always parallel to **E**. In general, however, **D** and **E** are in different directions.

The quantities ε_x, ε_y, and ε_z are known as the *principal dielectric permittivities* of the medium, and

$$n_x = \sqrt{\frac{\varepsilon_x}{\varepsilon_0}},\quad n_y = \sqrt{\frac{\varepsilon_y}{\varepsilon_0}},\quad \text{and } n_z = \sqrt{\frac{\varepsilon_z}{\varepsilon_0}} \tag{4.5}$$

are said to be the *principal refractive indices* of the medium. In the preceding equation, ε_0 represents the dielectric permittivity of free space. Depending on the relative values of ε_x, ε_y, and ε_z, the medium is classified as isotropic, uniaxial, or biaxial, as described here:

Isotropic: $$\varepsilon_x = \varepsilon_y = \varepsilon_z. \tag{4.6}$$

Uniaxial: $$\varepsilon_x = \varepsilon_y \neq \varepsilon_z. \tag{4.7}$$

Biaxial: $$\varepsilon_x \neq \varepsilon_y \neq \varepsilon_z. \tag{4.8}$$

As mentioned earlier, an anisotropic medium supports two rays known as the o ray and the e ray. The velocity of the o ray is the same in all directions, while the velocity of the e ray is different in different directions. Along a particular direction (fixed in the crystal), the velocities of the ordinary and extraordinary rays are equal; this direction is known as the *optic axis* of the crystal. In a uniaxial medium such as calcite, there is only one optic axis, which explains why such crystals are known as uniaxial crystals. In the general case of a biaxial crystal, there are two optic axes; i.e., there are two directions along which the two rays have the same velocity.

For a uniaxial medium, since $\varepsilon_x = \varepsilon_y$, the principal *x* and *y* directions can be arbitrarily chosen as long as they are perpendicular to the *z* axis. Furthermore, the quantities

$$n_o = n_x = n_y = \sqrt{\frac{\varepsilon_x}{\varepsilon_0}} = \sqrt{\frac{\varepsilon_y}{\varepsilon_0}} \quad \text{and} \quad n_e = n_z = \sqrt{\frac{\varepsilon_z}{\varepsilon_0}} \tag{4.9}$$

are known as ordinary and extraordinary refractive indices of the medium. Depending on the values of n_o and n_e, the uniaxial medium is divided into the following two categories:

Positive crystal: $n_e > n_o$ (e.g., quartz)
Negative crystal: $n_e < n_o$ (e.g., calcite)

The values of $n_o$and n_e for some uniaxial crystals are given in Table 4.1.

4.3 Plane Wave Propagation in Anisotropic Media

For a charge-free dielectric, Maxwell's equations simplify to (see Sec.2.2)

$$\nabla \cdot \mathbf{D} = 0, \tag{4.10}$$

$$\nabla \cdot \mathbf{H} = 0, \tag{4.11}$$

Table 4.1 Ordinary and extraordinary refractive indices for some uniaxial crystals (adapted from Refs. 1 and 2).

Name of Crystal	**Wavelength**	n_o	n_e
Calcite	4046 Å 5890 Å 7065 Å	1.68134 1.65835 1.65207	1.49694 1.48640 1.48359
Quartz	5890 Å	1.54424	1.55335
Lithium niobate	6000 Å	2.2967	2.2082
Potassium dihydrogen phosphate (KDP)	6328 Å	1.50737	1.46685
Ammonium dihydrogen phosphate (ADP)	6328 Å	1.52166	1.47685

$$\nabla \times \mathbf{E} = -\mu \frac{\partial \mathbf{H}}{\partial t}, \tag{4.12}$$

and

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t}, \tag{4.13}$$

where (as in Chapter 2) we have assumed that $\mathbf{B} = \mu\,\mathbf{H}$, and for most of the dielectrics, we may assume that $\mu \approx \mu_0 = 4\pi \times 10^{-7}$ N s^2/C^2, μ_0 being the magnetic permeability of vacuum.

For plane waves propagating in the direction of **k**, we may write

$$\mathbf{D} = \mathbf{D}_0\, e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}, \tag{4.14}$$

$$\mathbf{E} = \mathbf{E}_0\, e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}, \tag{4.15}$$

and

$$\mathbf{H} = \mathbf{H}_0\, e^{i(\omega t - \mathbf{k}\cdot\mathbf{r})}, \tag{4.16}$$

where $\mathbf{D}_0$, $\mathbf{E}_0$, and $\mathbf{H}_0$ are space- and time-independent vectors but may, in general, be complex. The equations $\nabla \cdot \mathbf{D} = 0$ and $\nabla \cdot \mathbf{H} = 0$ readily give

$$\mathbf{k} \cdot \mathbf{D} = 0, \tag{4.17}$$

and

$$\mathbf{k} \cdot \mathbf{H} = 0. \tag{4.18}$$

The two preceding equations imply that **D** and **H** are perpendicular to **k**. If we substitute the plane wave solutions in Maxwell's Eqs. (4.12) and (4.13), we obtain (see Section 2.3)

$$\mathbf{H} = \frac{1}{\omega\mu_0}(\mathbf{k} \times \mathbf{E}) \tag{4.19}$$

and

$$\mathbf{D} = -\frac{1}{\omega}(\mathbf{k} \times \mathbf{H}). \tag{4.20}$$

Thus, **D**, **E**, and **k** lie in a plane that is perpendicular to **H**. The phase front propagates along **k** with a wave velocity (also known as the *phase velocity*) given by

$$v_{\mathrm{w}} = \frac{\omega}{k} = \frac{c}{n_{\mathrm{w}}}, \tag{4.21}$$

where $k = \mathbf{k}/\hat{\boldsymbol{\kappa}}$, with $\hat{\boldsymbol{\kappa}}$ being the unit vector along **k**. Thus, $k = (\omega/c)\, n_{\mathrm{w}}$, where n_{w} is known as the *wave refractive index*. It should be noted that the direction of energy propagation is along the Poynting vector **S** = **E** × **H**, which is, in general, different from that of **k**.

4.3.1 Polarization eigenmodes

In the following discussion, we obtain the directions of the **D** vector and the corresponding values of n_{w} for the two polarization eigenmodes, i.e., the o wave and the e wave. Substituting **H** from Eq. (4.19) into Eq. (4.20), we obtain

$$\mathbf{D} = \frac{1}{\omega^2 \mu_0}\, [(\mathbf{k}\cdot\mathbf{k})\mathbf{E} \;-\; (\mathbf{k}\cdot\mathbf{E})\, \mathbf{k}], \tag{4.22}$$

where we have used the vector identity

$$\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}\,(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}\,(\mathbf{A}\cdot\mathbf{B}).$$

Further, using $\mathbf{k} = k\,\hat{\boldsymbol{\kappa}}$ and $k^2 = \dfrac{\omega^2}{c^2} n_{\mathrm{w}}^2 = \omega^2 \varepsilon_0 \mu_0 n_{\mathrm{w}}^2$ in Eq. (4.22), one obtains

$$\mathbf{D} = n_{\mathrm{w}}^2 \varepsilon_0\, [\mathbf{E} \;-\; (\hat{\boldsymbol{\kappa}}\cdot\mathbf{E})\, \hat{\boldsymbol{\kappa}}]\,. \tag{4.23}$$

Using the principal axis coordinate system, we write $D_x = \varepsilon_0 n_x^2 E_x$, $D_y = \varepsilon_0 n_y^2 E_y$, and $D_z = \varepsilon_0 n_z^2 E_z$. The *x*, *y*, and *z* components of Eq. (4.23) then give the following three homogeneous equations:

$$\left(\frac{n_x^2}{n_{\mathrm{w}}^2} + \kappa_x^2 - 1\right) E_x + \kappa_x \kappa_y E_y + \kappa_x \kappa_z E_z = 0, \tag{4.24}$$

$$\kappa_y \kappa_x E_x + \left(\frac{n_y^2}{n_{\mathrm{w}}^2} + \kappa_y^2 - 1\right) E_y + \kappa_y \kappa_z E_z = 0, \tag{4.25}$$

and

$$\kappa_z \kappa_x E_x + \kappa_z \kappa_y E_y + \left(\frac{n_z^2}{n_{\mathrm{w}}^2} + \kappa_z^2 - 1\right) E_z = 0, \tag{4.26}$$

where κ_x, κ_y, and κ_z are the *x*, *y*, *z* components of the unit vector $\hat{\boldsymbol{\kappa}}$, such that $\kappa_x^2 + \kappa_y^2 + \kappa_z^2 = 1$.

For nontrivial solutions of Eqs. (4.24)–(4.26), we must have

$$\begin{vmatrix} \frac{n_x^2}{n_w^2} + \kappa_x^2 - 1 & \kappa_x \kappa_y & \kappa_x \kappa_z \\ \kappa_y \kappa_x & \frac{n_y^2}{n_w^2} + \kappa_y^2 - 1 & \kappa_y \kappa_z \\ \kappa_z \kappa_x & \kappa_z \kappa_y & \frac{n_z^2}{n_w^2} + \kappa_z^2 - 1 \end{vmatrix} = 0. \qquad (4.27)$$

For a given direction of propagation (i.e., for given values of κ_x, κ_y, and κ_z), the solution of the preceding equation gives us two possible values of n_w. It appears that Eq. (4.27) is a cubic equation in n_w^2 and that there should be three roots of n_w^2; however, the coefficient of n_w^6 will always be zero, and hence, there will be only two roots.

4.3.1.1 Wave propagation along principal axes

It is extremely easy to find the roots of Eq. (4.27) when the direction of propagation is along one of the principal axes. Let the direction of wave propagation be along the *x* axis. In this case, $\kappa_x = 1$, $\kappa_y = 0$, $\kappa_z = 0$, and Eq. (4.27) becomes

$$\frac{n_x^2}{n_w^2}\left(\frac{n_y^2}{n_w^2} - 1\right)\left(\frac{n_z^2}{n_w^2} - 1\right) = 0, \qquad (4.28)$$

which gives the two permissible values of n_w as

$$n_{w1} = n_y \text{ and } n_{w2} = n_z. \qquad (4.29)$$

Substituting $n_{w1} = n_y$ along with $\kappa_x = 1$ and $\kappa_y = \kappa_z = 0$ in Eqs. (4.24) and (4.26), we obtain $E_x = 0$ and $E_z = 0$. This means that for $n_w = n_y$, the electric field **E** and, hence, **D** is along the *y* direction.

Similarly, it can be shown that for the second root $n_w = n_z$, the electric field **E** and **D** are along the *z* direction. Thus, if the wave is propagating along the *x* axis, the electric field **E** and, hence, **D** of the two eigenmodes are polarized along the *y* and *z* axes, and the corresponding refractive indices are n_y and n_z, respectively. Similarly, one can show that

(i) If the wave is propagating along the *y* axis, **E** and **D** of the two eigenmodes are polarized along the *z* and *x* axes, and the corresponding refractive indices are n_z and n_x and, respectively.

(ii) If the wave is propagating along the *z* axis, **E** and **D** of the two eigenmodes are polarized along the *x* and *y* axes, and the corresponding refractive indices are n_x and n_y, respectively. It may be noted that for a uniaxial medium, $n_x = n_y$ and hence, the wave velocities of both eigenmodes (o-wave and e-wave) will be equal. Thus, the *z* axis will be the optical axis of the medium.

If the direction of the propagation is not along any of the principal axes, then the roots of Eq. (4.27) are difficult to obtain for the general case of a biaxial crystal. However, for a uniaxial medium, the roots are easy to obtain and are discussed in the following section.

4.3.1.2 Wave propagation in a uniaxial medium: arbitrary direction

Let the direction of propagation be such that the **k** vector makes an angle ψ with the *z* axis. In a uniaxial medium, the *x* and *y* principal axes can be chosen arbitrarily in the plane perpendicular to the *z* axis. We choose the *x* axis lying in the **k**-*z* plane; obviously, the *y* axis will be perpendicular to the **k**-*z* plane (see Fig. 4.1). Thus, we may write

$$\kappa_x = \sin\psi,\ \kappa_y = 0, \text{ and } \kappa_z = \cos\psi.$$

Equations (4.24) through (4.26), therefore, become

$$\left(\frac{n_o^2}{n_w^2} - \cos^2\psi\right) E_x + \left(\sin\psi\,\cos\psi\right) E_z = 0\,, \tag{4.30}$$

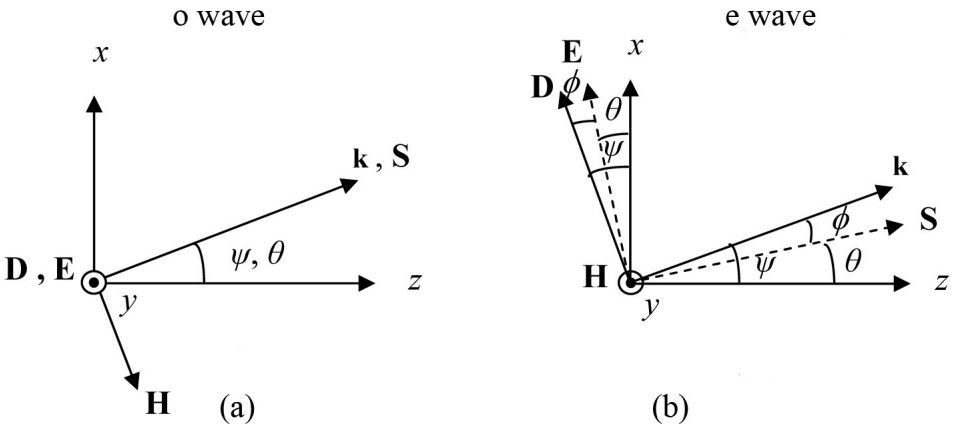


Figure 4.1 Relative directions of various vectors with respect to the optic axis in a uniaxial crystal for (a) ordinary and (b) extraordinary waves.

$$\left(\frac{n_{\mathrm{o}}^2}{n_{\mathrm{w}}^2} - 1\right) E_y = 0\,, \tag{4.31}$$

and

$$\sin\psi\,\cos\psi\,E_x + \left(\frac{n_{\mathrm{e}}^2}{n_{\mathrm{w}}^2} - \sin^2\psi\right) E_z = 0\,, \tag{4.32}$$

where we have used the fact that for a uniaxial medium, $n_x = n_y = n_{\mathrm{o}}$ and $n_z = n_{\mathrm{e}}$. Once again, we have a set of three homogenous equations and, for nontrivial solutions, the determinant must be zero; i.e.,

$$\begin{vmatrix} \dfrac{n_{\mathrm{o}}^2}{n_{\mathrm{w}}^2} - \cos^2\psi & 0 & \sin\psi\,\cos\psi \\ 0 & \dfrac{n_{\mathrm{o}}^2}{n_{\mathrm{w}}^2} - 1 & 0 \\ \sin\psi\,\cos\psi & 0 & \dfrac{n_{\mathrm{e}}^2}{n_{\mathrm{w}}^2} - \sin^2\psi \end{vmatrix} = 0\,,$$

i.e.,

$$\left(\frac{n_{\mathrm{o}}^2}{n_{\mathrm{w}}^2} - 1\right)\left[\left(\frac{n_{\mathrm{o}}^2}{n_{\mathrm{w}}^2} - \cos^2\psi\right)\left(\frac{n_{\mathrm{e}}^2}{n_{\mathrm{w}}^2} - \sin^2\psi\right) - \sin^2\psi\,\cos^2\psi\right] = 0\,, \tag{4.33}$$

which gives the following two solutions for n_{w}:

Solution 1: $n_{\mathrm{w}} = n_{\mathrm{o}}$. The corresponding wave velocity is

$$v_{\mathrm{w}} = v_{\mathrm{wo}} = \frac{c}{n_{\mathrm{o}}}. \quad \text{(o wave)} \tag{4.34}$$

Since the wave velocity is independent of the direction of the wave, it is referred to as the ordinary wave (o wave). Substituting $n_{\mathrm{w}} = n_{\mathrm{o}}$ in Eqs. (4.30) and (4.32), we obtain $E_x = 0 = E_z$. Thus, for the o wave, the **D** vector (and the **E** vector) is *y* polarized, i.e., perpendicular to the plane containing the **k** vector and the optic axis. The relative directions of various vectors (**D**, **E**, **k**, **S**, and **H**) with respect to the optic axis for the o wave are shown in Fig. 4.1(a).

Solution 2: The second solution of Eq. (4.33) is obtained from the equation

$$\left(\frac{n_o^2}{n_w^2} - \cos^2\psi\right)\left(\frac{n_e^2}{n_w^2} - \sin^2\psi\right) - \sin^2\psi\cos^2\psi = 0,$$

resulting in
$$\frac{1}{n_w^2} = \frac{1}{n_{we}^2} = \frac{\cos^2\psi}{n_o^2} + \frac{\sin^2\psi}{n_e^2}. \tag{4.35}$$

The corresponding wave velocity is given by

$$v_{we}^2 = \frac{c^2}{n_{we}^2} = \frac{c^2}{n_o^2}\cos^2\psi + \frac{c^2}{n_e^2}\sin^2\psi. \text{ (e wave)} \tag{4.36}$$

Since the wave velocity is dependent on the direction of the wave, it is referred to as the extraordinary wave (e wave) and, hence, the subscript e.

In order to obtain the directions of the **D** and **E** vectors of the e wave, we substitute $n_w = n_{we}$ from Eq. (4.35) in Eqs. (4.30) to (4.32), giving

$$E_y = 0$$

and

$$\frac{n_e^2 E_z}{n_o^2 E_x} = \frac{D_z}{D_x} = -\tan\psi. \tag{4.37}$$

The preceding equation implies that the **D** vector of the e wave makes an angle ψ with the x axis and is perpendicular to the **k** vector [see Fig. 4.1(b)]. Further, **D** is not parallel to **E**; the direction of **E** can be obtained in terms of ψ, as discussed in the following:

If θ is the angle made by the **E** vector with the x axis, then

$$\tan\theta = -\frac{E_z}{E_x} = -\frac{\varepsilon_x D_z}{\varepsilon_z D_x} = -\frac{n_o^2}{n_e^2}\left(\frac{D_z}{D_x}\right),$$

i.e.,
$$\tan\theta = \frac{n_o^2}{n_e^2}\tan\psi. \tag{4.38}$$

Since **S** is perpendicular to **E** and **H**, it will also lie in the **k**-z plane and will make an angle θ with z (the optic axis). Thus, for the e wave, the directions of wave propagation and ray propagation are given by angles ψ and θ, respectively, which are related to each other through Eq. (4.38). It is clear from Eq. (4.38) that $\theta < \psi$ for a positive crystal ($n_e > n_o$), and $\theta > \psi$ for a negative crystal ($n_e < n_o$).

The relative directions of various vectors (**D**, **E**, **k**, **S**, and **H**) with respect to the optic axis for the e wave for a positive uniaxial crystal are shown in Fig. 4.1(b). The criteria for identifying the **D** vectors of the o wave and the e wave is as follows:

(i) For both the o wave and the e wave, **D** is always perpendicular to **k** (i.e., $\mathbf{D} \cdot \mathbf{k} = 0$).
(ii) For the o wave, **D** is perpendicular to the **k**-*z* plane.
(iii) For the e wave, **D** lies in the **k**-*z* plane.

It should be noted that for an e ray, the direction of the wave propagation (i.e., of **k**) is different from the direction of the ray propagation (i.e., of **S**). This leads to different values of the wave velocity (with which the phase front moves) and the ray velocity (with which the energy moves), as discussed in the following section.

4.4 Ray Velocity and Ray Refractive Index

The direction of ray propagation (or energy propagation) is along the Poynting vector **S**, which is given by

$$\mathbf{S} = \mathbf{E} \times \mathbf{H}. \tag{4.39}$$

The ray velocity (or the energy transmission velocity) v_r is defined as

$$v_r = \frac{S}{u}, \tag{4.40}$$

where u is the energy density given by[2]

$$u = \frac{1}{2}(\mathbf{D}\cdot\mathbf{E} + \mathbf{B}\cdot\mathbf{H}) = \frac{1}{2}(\mathbf{D}\cdot\mathbf{E} + \mu_0 \mathbf{H}\cdot\mathbf{H}). \tag{4.41}$$

Substituting for **H** and **D** from Eqs. (4.19) and (4.20), we obtain

$$u = \frac{1}{2\omega}[(\mathbf{H}\times\mathbf{k})\cdot\mathbf{E} + (\mathbf{k}\times\mathbf{E})\cdot\mathbf{H}]$$

$$= \frac{1}{2\omega}[\mathbf{k}\cdot(\mathbf{E}\times\mathbf{H}) + \mathbf{k}\cdot(\mathbf{E}\times\mathbf{H})] = \frac{1}{\omega}\mathbf{k}\cdot\mathbf{S}. \tag{4.42}$$

Thus, Eq. (4.40) becomes

$$v_r = \frac{\omega S}{\mathbf{k}\cdot\mathbf{S}} = \frac{\omega}{k\cos\phi} = \frac{v_w}{\cos\phi},$$

where ϕ is the angle between **k** and **S**. Further, since **D** is perpendicular to **k**, and **E** is perpendicular to **S,** ϕ is also the angle between **D** and **E** [see Fig. 4.1(b)]. The wave and ray velocities are therefore related as

$$v_w = v_r \cos\phi . \tag{4.43}$$

The ray refractive index n_r is defined as

$$n_r = \frac{c}{v_r} = \frac{c}{v_w}\cos\phi = n_w \cos\phi . \tag{4.44}$$

For the o component, **D** and **E** are parallel, implying that $\phi = 0$; therefore,

$$v_{ro} = v_{wo} = \frac{c}{n_o}, \text{ and } n_{ro} = n_{wo} = n_o . \tag{4.45}$$

On the other hand, for the e component, **D** and **E** are not parallel; therefore,

$$v_{re} \neq v_{wo}, \text{ and } n_{re} \neq n_{wo} . \tag{4.46}$$

The ray velocity and the ray refractive index for the e ray can also be expressed in terms of the angle θ (see Example 4.1) as

$$\frac{1}{v_{re}^2} = \frac{\cos^2\theta}{c^2/n_o^2} + \frac{\sin^2\theta}{c^2/n_e^2} \tag{4.47}$$

and

$$n_{re}^2(\theta) = n_o^2 \cos^2\theta + n_e^2 \sin^2\theta . \tag{4.48}$$

It is clear from Eqs. (4.45) and (4.47) that the velocity of the o ray v_{ro} is the same in all directions, whereas the velocity of the e ray v_{re} is θ dependent. In a special case, v_{re} is equal to c/n_o when the ray propagates parallel to the optic axis (θ = 0), and v_{re} is equal to c/n_e when the ray propagates perpendicular to the optic axis ($\theta = \pi/2$).

Using the preceding analysis, one can obtain the direction of ray propagation in a uniaxial medium if the direction of wave propagation is known, or vice versa. However, we often know the direction of a light beam incident on an anisotropic medium and need to know the directions of the refracted rays. The directions of the refracted rays can be obtained using Huygens' construction, which further uses ray surfaces. In the following section we discuss the ray

surfaces, and in Section 4.6 we include ray surfaces in the discussion of the phenomenon of refraction in a uniaxial medium.

4.4.1 Ray surfaces

The ray velocities of the o ray and e ray are given by Eqs. (4.45) and (4.47), respectively. If we plot v_{ro} and v_{re} as a function of θ, we will obtain a sphere and an ellipsoid of revolution (see Fig. 4.2), respectively, which are known as the *ray surfaces*. These are the surfaces on which the different o rays and e rays, starting from a point source sitting at the origin, would end up after a unit time. The sphere and the ellipsoid will be touching each other along the optic axis ($\theta = 0$) because both rays travel with same velocity c/n_o. In a negative crystal ($n_e < n_o$), since $v_{re} > v_{ro}$, the ellipsoid of revolution will lie outside the sphere, as shown in Fig. 4.2(a). For a positive crystal ($n_e > n_o$), since $v_{re} < v_{ro}$, the ellipsoid of revolution will lie inside the sphere, as shown in Fig. 4.2(b).

Example 4.1 Obtain expressions for the ray velocity and ray refractive indices of the e ray in terms of the ray propagation angle θ.

Solution: We note that $\phi = \psi - \theta$ [Fig. 4.1(b)] and use Eq. (4.38) to express ϕ and v_{we} in terms of θ as

$$\tan\phi = \tan(\psi - \theta) = \frac{\tan\psi - \tan\theta}{1 + \tan\psi\tan\theta} = -\frac{\left(1 - \dfrac{n_e^2}{n_o^2}\right)\tan\theta}{1 + \dfrac{n_e^2}{n_o^2}\tan^2\theta},$$

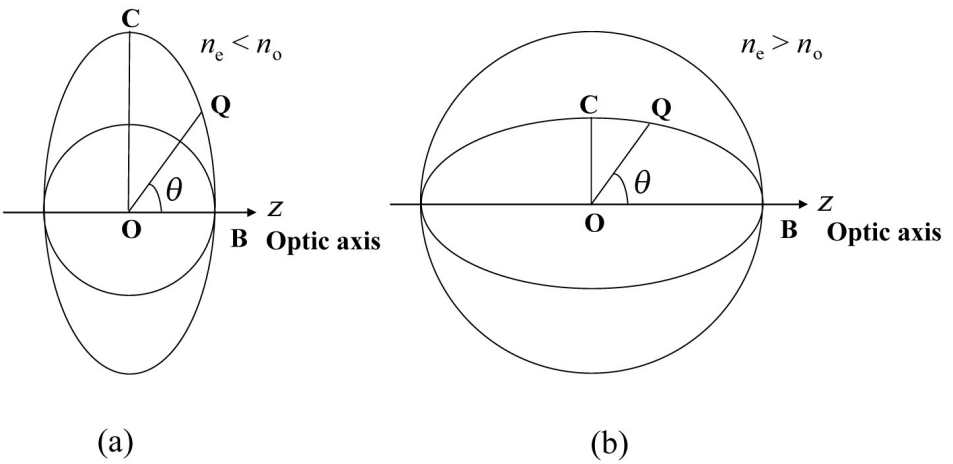


Figure 4.2 Schematic of ray surfaces for (a) negative and (b) positive crystals.

i.e., $$\sec^2\phi = 1+\tan^2\phi = \sec^2\theta \frac{1+\frac{n_e^4}{n_o^4}\tan^2\theta}{\left(1+\frac{n_e^2}{n_o^2}\tan^2\theta\right)^2}. \tag{4.49}$$

Further, using Eq. (4.36),

$$v_{we}^2 = \frac{c^2}{n_o^2}\cos^2\psi\left(1+\frac{n_o^2}{n_e^2}\tan^2\psi\right)$$

$$= \frac{c^2}{n_o^2}\frac{1+\frac{n_o^2}{n_e^2}\tan^2\psi}{1+\tan^2\psi} = \frac{c^2}{n_o^2}\frac{1+\frac{n_e^2}{n_o^2}\tan^2\theta}{1+\frac{n_e^4}{n_o^4}\tan^2\theta}. \tag{4.50}$$

Now, from Eq. (4.43),

$$v_{re}^2 = v_{we}^2\sec^2\phi.$$

Substituting $\sec^2\phi$ from Eq. (4.49) and v_{we}^2 from Eq. (4.50) into the preceding equation, we obtain

$$v_{re}^2 = \frac{c^2}{n_o^2}\frac{\sec^2\theta}{1+\frac{n_e^2}{n_o^2}\tan^2\theta} = \frac{c^2}{n_o^2\cos^2\theta + n_e^2\sin^2\theta},$$

which can be expressed as

$$\frac{1}{v_{re}^2} = \frac{\cos^2\theta}{c^2/n_o^2} + \frac{\sin^2\theta}{c^2/n_e^2}.$$

The corresponding ray refractive index is thus given by

$$n_{re}^2(\theta) = \frac{c^2}{v_{re}^2} = n_o^2\cos^2\theta + n_e^2\sin^2\theta.$$

In the preceding sections, we have discussed the characteristics of the two polarization eigenmodes (o wave and e wave) in a uniaxial medium. In the case

of a biaxial medium, the mathematical approach to obtain these eigenmodes is a bit more difficult. In the following section, we discuss a geometrical construction known as *index ellipsoid*, which can be used to obtain these states and the corresponding refractive indices.

4.5 Index Ellipsoid

The index ellipsoid is described by the following equation:

$$\frac{x^2}{n_x^2}+\frac{y^2}{n_y^2}+\frac{z^2}{n_z^2}=1. \tag{4.51}$$

This equation represents a generalized ellipsoid with its major axes along the x, y, and z directions with lengths OA = n_x, OB = n_y, and OC = n_z, respectively (Fig. 4.3). The directions of the **D** vector for the two polarization eigenmodes are determined from the index ellipsoid as follows:

(i) We first draw a plane perpendicular to the direction of the wave propagation **k** passing through the origin. This plane will intersect the index ellipsoid in general along an ellipse.

(ii) The directions of the major and minor axes of this ellipse give the directions of the **D** vector for the two polarization eigenmodes, and the lengths of the semi-axes give the corresponding wave refractive indices.

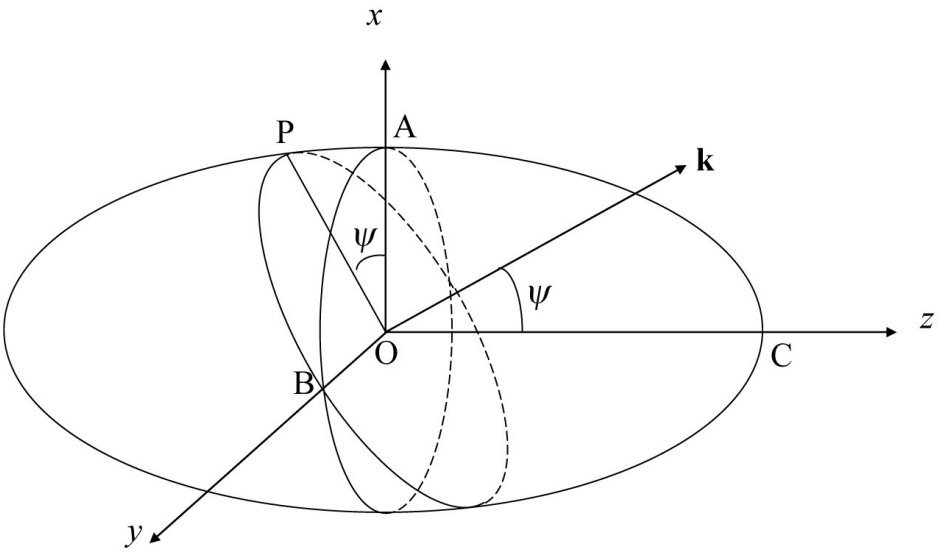


Figure 4.3 Index ellipsoid described by Eq. (4.51).

It can be shown that for a biaxial crystal, there exist two directions of propagation for which the intersection of the ellipsoid and the plane perpendicular to **k** will be a circle, implying that both of the polarizations travel with the same refractive indices. These directions of **k** are the two optic axes of the medium. In the case of a uniaxial medium, $n_x = n_y \neq n_z$, and there is only one such direction (optic axis) that is also parallel to the *z* axis.

4.5.1 Uniaxial medium

In order to demonstrate the use of the index ellipsoid, we consider a uniaxial medium with $n_x = n_y = n_o$ and $n_z = n_e$. In this case, the index ellipsoid is an ellipsoid of revolution around the *z* axis and is described by the following equation:

$$\frac{x^2}{n_o^2} + \frac{y^2}{n_o^2} + \frac{z^2}{n_e^2} = 1. \tag{4.52}$$

It is clear that the *z* axis is the optic axis, as the central plane perpendicular to the *z* axis will cut the ellipsoid in a circle. Let the direction of wave propagation **k** make an angle ψ with the *z* axis. Without any loss of generality, we can choose the *x* axis lying in the **k**-*z* plane. Now, we draw a plane perpendicular to **k** and passing through O that cuts the ellipsoid in an ellipse with one of its axes parallel to the *y* axis and with the semi-axis OB = n_o. The other axis lies in the *x*-*z* plane, with semi-axis OP = $n_{we}(\psi)$, which can be determined as follows: The point P lies on an ellipse whose equation is

$$\frac{x^2}{n_o^2} + \frac{z^2}{n_e^2} = 1. \tag{4.53}$$

The (*x*, *z*) coordinates of point P are $x = \mathrm{OP}\cos\psi$ and $z = -\mathrm{OP}\sin\psi$, with $\mathrm{OP} = n_{we}(\psi)$. Substituting these values of *x* and *z* in the preceding equation, we obtain

$$\frac{1}{n_{we}^2(\psi)} = \frac{\cos^2\psi}{n_o^2} + \frac{\sin^2\psi}{n_e^2}. \tag{4.54}$$

The preceding equation is the same as Eq. (4.35). Thus, the **D** vector of the ordinary component is along the *y* axis, with wave refractive index n_o, and the **D** vector of the extraordinary component is along OP, with wave refractive index $n_{we}(\psi)$.

4.6 Refraction in a Uniaxial Medium

In this section, we will obtain the directions of the refracted o ray and e ray and their **D** vectors in a uniaxial medium, corresponding to an incident plane electromagnetic wave. We will consider a negative crystal, in this case, calcite; a similar analysis can also be carried out for positive crystals.

4.6.1 Normal incidence

We first assume a plane wave incident normally on a uniaxial crystal, as shown in Fig. 4.4. Without any loss of generality, we can always choose the optic axis to lie in the plane of the paper, which is shown as a dashed line in Fig. 4.4. Let BD represent the incident wavefront. In order to determine the o ray, with the point B as the center, we draw a sphere of radius c/n_o. Similarly, we draw another sphere (of the same radius) from the point D. The common tangent plane to these spheres is shown as OO′, which represents the wavefront corresponding to the ordinary refracted ray. If we join point B to the point of contact O, then corresponding to the incident ray AB, the direction of the refracted o ray will be along BO.

In order to determine the refracted e ray, we draw an ellipse (centered at point B) with its minor axis (equal to c/n_o) along the optic axis and with the major axis equal to c/n_e. The ellipsoid of revolution is obtained by rotating the ellipse about the optic axis. Similarly, we draw another ellipsoid of revolution centered at point D. The common tangent plane to these ellipsoids (which will be perpendicular to **k**) is shown as EE′ in Fig. 4.4. If we join point B to the point of contact E (between the ellipsoid of revolution and the tangent plane EE′), corresponding to the incident ray AB, the direction of the refracted e ray will be along BE.

We should mention here that the direction of **k** is the *same* for both o and e waves, i.e., both are along BO. However, corresponding to incident ray AB, the o ray will propagate along BO, and the extraordinary ray will propagate in a different direction (along BE) [Fig. 4.4(a)]. Thus, an incident ray will split into two rays propagating in different directions, and when these rays leave the crystal, we have two linearly polarized beams.

If we change the direction of the optic axis [Fig, 4.4(b)], although the direction of the o ray remains the same, the e ray will propagate in a different direction. Thus, if a ray is incident normally on a calcite crystal, and if the crystal is rotated about the normal, then the optic axis and the e ray will also rotate (about the normal) on the periphery of a cone, lying in the plane containing the normal and the optic axis [see Fig. 4.4(b)].

As mentioned in Section 4.3.1.2, the directions of the **D** vibrations for the o ray are normal to the optic axis and the vector **k**; as such, the directions of these vibrations in this case will be normal to the plane of the paper and are shown as dots in Fig. 4.4. Similarly, since the direction of vibrations for the e ray is perpendicular to **k** and lies in the plane containing the e ray and the optic axis, the vibrations are along the short, straight lines drawn on the e ray in Fig. 4.4.

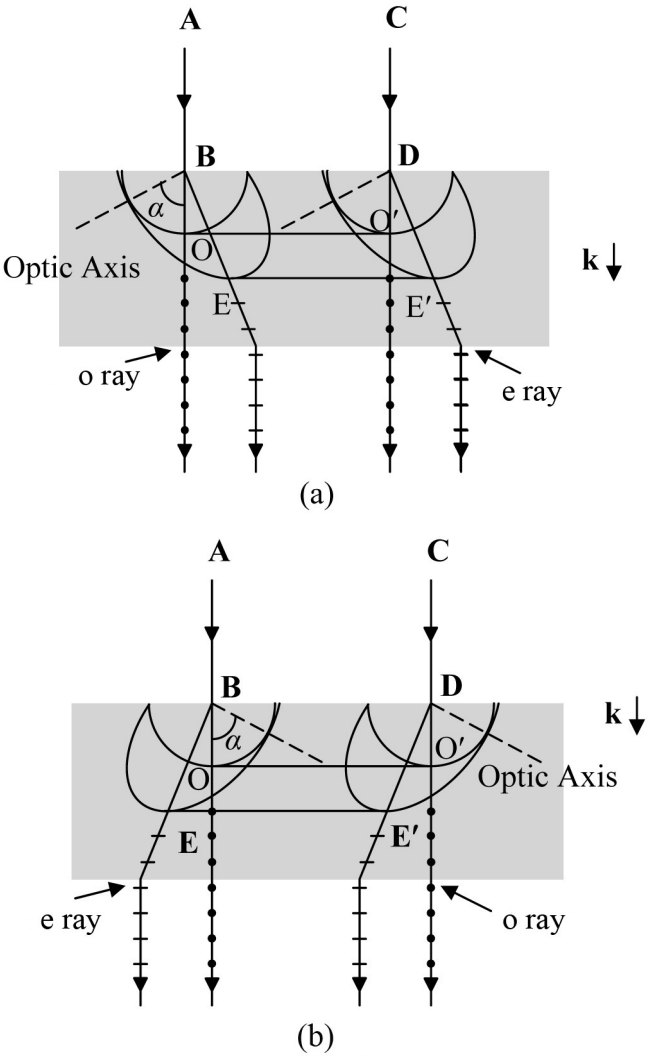


Figure 4.4 The refraction of a plane wave incident on a negative crystal whose optic axis is along the dashed line. The dots and the small lines show the directions of the **D** vector for the o ray and e ray, respectively. The crystals shown in (a) and (b) are rotated with respect to each other by an angle of 180 deg about the direction of the incident wave.

In the preceding case, we assumed that the optic axis makes an arbitrary angle α with the normal to the surface. In the special cases of $\alpha = \pi/2$ and $\alpha = 0$, the o ray and the e ray travel along the same directions, as shown in Figs. 4.5(a) through 4.5(c). It may be mentioned that Figs. 4.5(a) and 4.5(b) correspond to the same configuration, representing two different cross sections of the same set of spherical and ellipsoidal wavefronts. Figure 4.5(b) corresponds to the case in which the optic axis is normal to the plane of the paper, and as such, the section of the extraordinary wavefront in the plane of the paper will be a circle. Once again, both the o ray and the e ray travel along the same direction.

Corresponding to Figs. 4.5(a) and 4.5(b), if the incident wave is polarized perpendicular to the optic axis, it will propagate as an o wave with velocity c/n_o. On the other hand, if the incident wave is polarized parallel to the optic axis, it will propagate as an e wave with velocity c/n_e. In Fig. 4.5(c), the optic axis is normal to the surface, and both waves will travel with the same velocity.

Notice that in the configuration shown in Figs. 4.5(a) and 4.5(b), although both waves travel in the same direction, they propagate with different velocities. This phenomenon is used in the fabrication of quarter- and half-wave plates (see Sec. 4.7.2). On the other hand, in the configuration shown in Fig. 4.5(c), both of the waves not only travel in the same direction but also propagate with the same velocity.

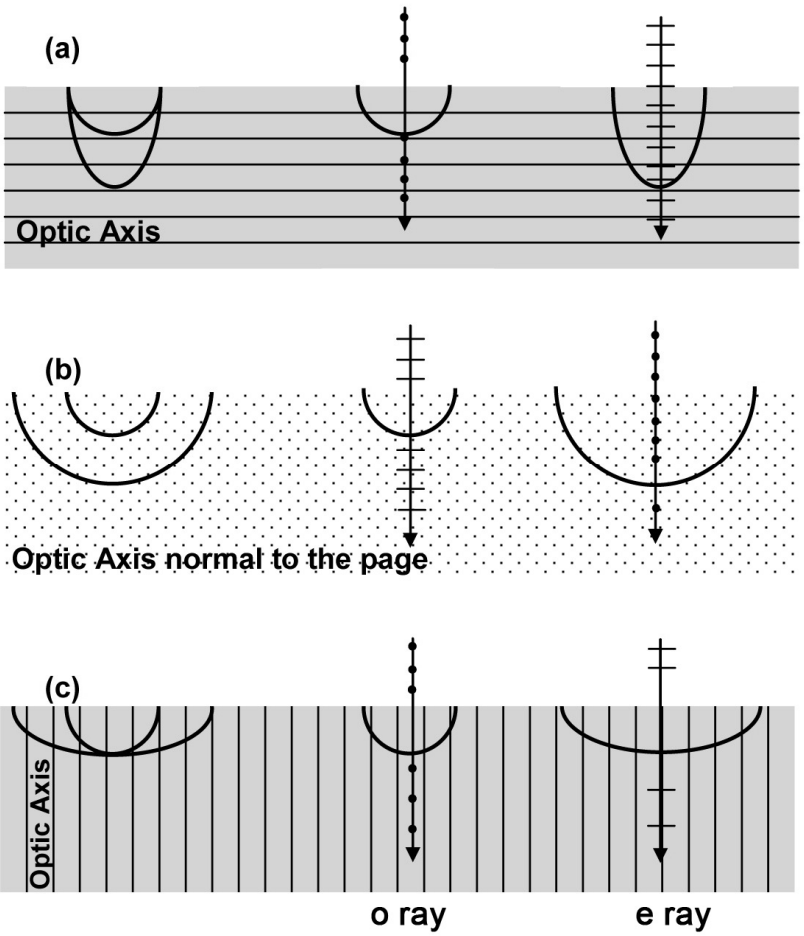


Figure 4.5 Propagation of a plane wave incident normally on a negative uniaxial crystal. In (a) and (c), the optic axis is shown as parallel straight lines, and in (b), the optic axis is perpendicular to the plane of the figure and is shown as dots. In each case, the e and o rays travel in the same direction.

4.6.2 Oblique incidence

We next consider the case of a plane wave incident obliquely on a negative uniaxial crystal [Fig. 4.6(a)]. Once again, we use Huygens' construction to determine the refracted wavefronts. Let BD represent the incident wavefront. If the time taken for the disturbance to reach point F from D is t, with B as the center, we draw a sphere of radius $(c/n_o)t$ and an ellipsoid of revolution of semi-minor and semi-major axes $(c/n_o)t$ and $(c/n_e)t$, respectively; the semi-minor axis is along the optic axis. From point F, we draw tangent planes FO and FE to the sphere and the ellipsoid of revolution, respectively. These planes represent the refracted wavefronts corresponding to the o ray and the e ray, respectively. If the points of contact are O and E, the ordinary and extraordinary refracted rays will propagate along BO and BE, respectively. The directions of **D** vibrations of these rays are shown as dots and short lines, respectively, and are obtained by using the general rules discussed in Section 4.3.1.2. The refracted wavefronts corresponding to the particular case of $\alpha = 0$ and $\alpha = \pi/2$ can be very easily obtained.

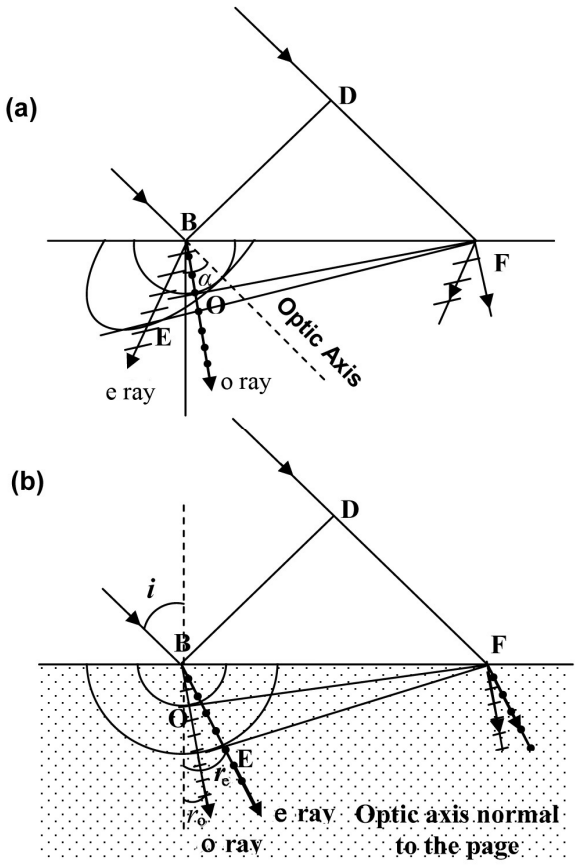


Figure 4.6 Refraction of a plane wave incident obliquely on a negative uniaxial crystal. (a) The direction of the optic axis is along the dashed line. (b) The optic axis is perpendicular to the plane of the paper.

Figure 4.6(b) corresponds to the case in which the optic axis is normal to the plane of incidence. The sections of both the wavefronts will be circles; consequently, the e ray will also satisfy Snell's law, and we will have

$$\frac{\sin i}{\sin r_e} = n_e . \tag{4.55a}$$

Of course, for the o ray, we will *always* have

$$\frac{\sin i}{\sin r_o} = n_o . \tag{4.55b}$$

4.7 Polarization Components Using Double Refraction

4.7.1 Linear polarizer

As discussed in Section 4.1, when an unpolarized beam enters an anisotropic crystal, it splits into two beams (i.e., ordinary and extraordinary), each of them being characterized by a certain linear SOP. If by some method one of the beams is eliminated, the transmitted beam will be linearly polarized, and the crystal will act as a linear polarizer. This is accomplished through total internal reflection in a Nicol prism, as discussed in the following section.

4.7.1.1 The Nicol prism

A Nicol prism consists of a calcite crystal that is cut into two prisms and cemented together with a layer of Canada balsam sandwiched between. The refractive index of Canada balsam (n = 1.55) lies between the ordinary and extraordinary refractive indices of calcite. An unpolarized beam entering the first prism is split into ordinary and extraordinary beams, as shown in Fig. 4.7. The angles of the prism are such that the o ray suffers total internal reflection at the calcite–Canada balsam interface, whereas the extraordinary component passes through the balsam, and the beam emerging from the crystal is linearly polarized.

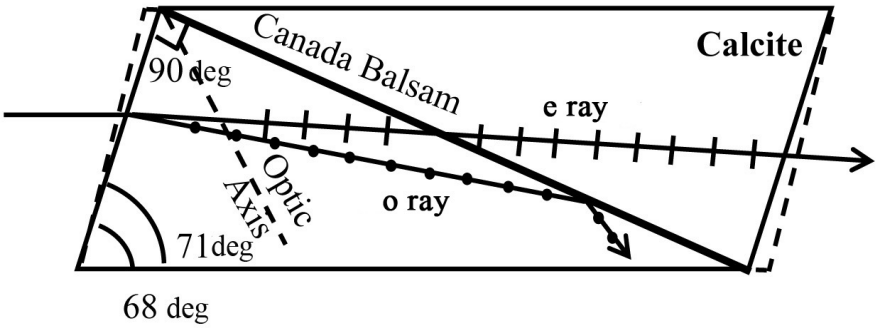


Figure 4.7 Schematic of Nicol prism and its working principle.

4.7.2 Quarter-wave plates and half-wave plates

Let us consider the normal incidence of a plane polarized beam on a calcite crystal whose optic axis is parallel to the surface of the crystal [Fig. 4.5 (a)]. We assume that the optic axis is along the *z* axis and that the *x* axis is the direction of propagation. If the incident beam is *y* polarized, the beam will propagate as an o wave, and the e wave will be absent. Similarly, if the incident beam is *z* polarized, the beam will propagate as an extraordinary wave, and the o wave will be absent. For any other SOP of the incident beam, both the ordinary and the extraordinary components will be present. Since $n_o \neq n_e$, the two beams will propagate with different phase velocities, and as such, when they come out of the crystal, they will not be in the same phase. Consequently, the emergent beam will be, in general, elliptically polarized.

If the thickness of the crystal plate is *d*, the phases of the ordinary and the extraordinary beams will change by $n_o kd$ and $n_e kd$, respectively, resulting in a phase difference δ given by

$$\delta = k\,d\,(n_o - n_e) = \frac{\omega}{c}(n_o - n_e)\,d\,. \tag{4.56}$$

If the thickness *d* of the crystal is such that $\delta = \pi/2$, the crystal plate is said to be a quarter-wave plate (QWP); a phase difference of $\pi/2$ implies a path difference of a quarter of a wavelength. The required thickness is given by

$$d = \frac{c}{\omega(n_o - n_e)}\frac{\pi}{2} = \frac{1}{4}\frac{\lambda_0}{(n_o - n_e)}. \tag{4.57}$$

On the other hand, if the thickness of the crystal plate is such that $\delta = \pi$, the crystal is said to be a half-wave plate (HWP). The required thickness is given by

$$d = \frac{\lambda_0}{2(n_o - n_e)}. \tag{4.58}$$

Clearly, if the thickness of the crystal plate is such that $\delta = 2\pi, 4\pi, \ldots$, the emergent beam will have the same SOP as the incident beam.

4.7.3 The Rochon prism

A Rochon prism is used to obtain two linear and orthogonal polarized components of a given light beam. It consists of two similar prisms (for instance, of calcite); the optic axis of the first prism is normal to the face of the prism, while the optic axis of the second prism is parallel to the edge, as shown in Fig. 4.8.

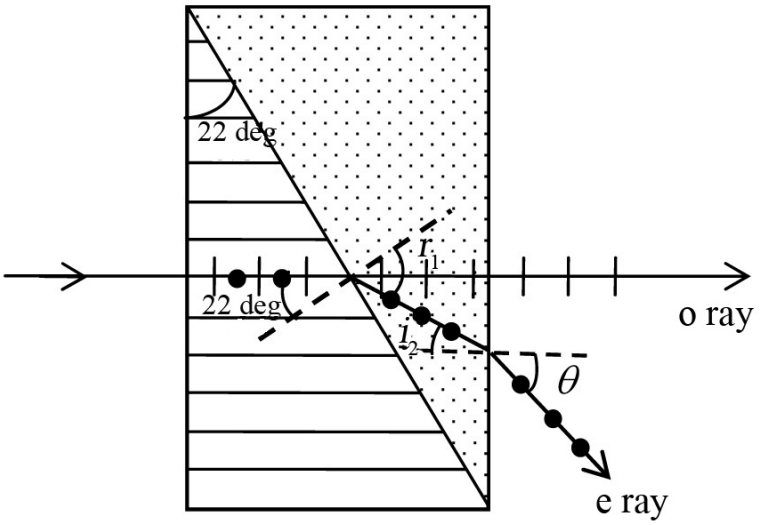


Figure 4.8 Production of two orthogonally polarized beams by a Rochon prism.

In the first prism, since the beam is propagating along the optic axis, both the o ray and the e ray will travel with the same velocity (c/n_o) and, hence, will see the same refractive index n_o. When the beam enters the second prism, the o ray (whose **D** is normal to the optic axis) will see the same refractive index n_o and go undeviated, as shown in Fig. 4.8. On the other hand, the e ray (whose **D** is along the optic axis) will see the refractive index n_e. For calcite, $n_e < n_o$, so the ray will bend away from the normal. Thus, the two orthogonally polarized components will come out in different directions. The splitting angle between them will be obtained as follows:

Let the prism angle be 22 deg. Since the optic axis is normal to the plane of incidence, the refracted e ray will obey Snell's law (see Fig. 4.8), and the angle of refraction will be determined from the following equation:

$$n_o \sin 22\,\text{deg} = n_e \sin r_1 .$$

Thus, at $\lambda_0 = 5890$ Å,

$$\sin r_1 = \frac{n_o}{n_e} \sin 22\,\text{deg} = \frac{1.6584}{1.4864} \times 0.375 \approx 0.418 \Rightarrow r_1 = 24.7\,\text{deg} .$$

Therefore, the angle of incidence at the second surface will be $i_2 = 24.7 - 22$ deg = 2.7 deg. The emerging angle will be given by

$$\sin\theta = n_e \sin\left(2.7\ \text{deg}\right) \approx 0.070 \Rightarrow \theta \approx 4\,\text{deg} .$$

The angular separation between the two output beams can be further increased by appropriately selecting the direction of the optic axis in the first prism, which is done in a Wollaston prism, as discussed in the next section.

4.7.4 The Wollaston prism

A Wollaston prism consists of two similar prisms (for instance, of calcite), with the optic axis of the first prism parallel to the face of the prism and the optic axis of the second prism parallel to the edge of the prism, as shown in Fig. 4.9.

Let us first consider the incidence of a *z*-polarized beam, as shown in Fig. 4.9(a). The beam will propagate as an o ray in the first prism and will see the refractive index n_o. When this beam enters the second prism, it will become an e ray and will see the refractive index n_e. For calcite, $n_e < n_o$, and therefore, the ray will bend away from the normal. The angle of refraction will be given by

$$n_o \sin 22 \text{ deg} = n_e \sin r_1.$$

We again assume the angle of the prism to be 22 deg, $n_o \approx 1.6584$, and $n_e \approx 1.4864$ (see Fig. 4.9). We readily obtain $r_1 \approx 24.7$ deg. Thus, the angle of incidence at the second surface will be $i_2 = 24.7 \text{ deg} - 22 \text{ deg} = 2.7$ deg. The output angle θ_1 will be given by $n_e \sin 2.7 \text{ deg} = \sin \theta_1 \Rightarrow \theta_1 \approx 4$ deg.

Next, we consider the incidence of a *y*-polarized beam, as shown in Fig. 4.9(b). The beam will propagate as an e ray in the first prism and as an o ray in the second prism. The angle of refraction will now be given by $n_e \sin 22 \text{ deg} = n_o \sin r_1' \Rightarrow r_1' \approx 19.62$ deg. Thus, the angle of incidence at the second interface will be $i_2' = 22 \text{ deg} - 19.62 \text{ deg} = 2.38$ deg. The output angle θ_2 will be given by $n_o \sin 2.38 \text{ deg} = \sin \theta_2 \Rightarrow \theta_2 \approx 3.95$ deg. Thus, if an unpolarized beam is incident on the Wollaston prism, the angular separation between the two orthogonally polarized output beams will be $\theta = (\theta_1 + \theta_2) \approx 7.95$ deg.

The birefringence discussed earlier is an example of linear birefringence, as the two polarization eigenmodes considered are linearly polarized. We also have examples where the two polarization eigenmodes are circularly or, in general, elliptically polarized. Accordingly, the medium is said to be circularly or elliptically birefringent (as mentioned in Section 3.4). In the following section, we discuss the propagation in a circularly birefringent medium.

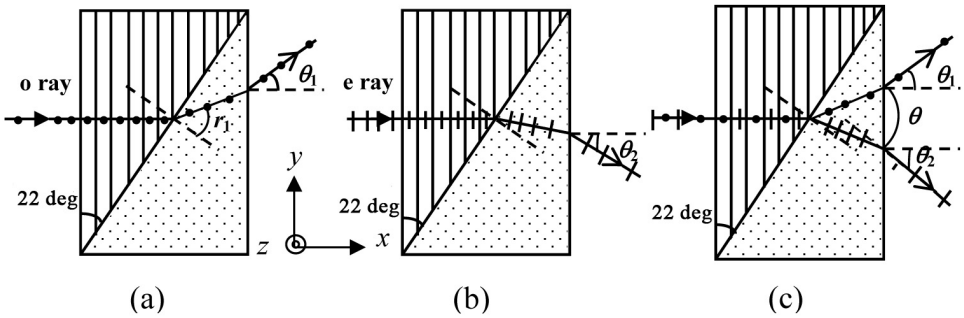


Figure 4.9 A Wollaston prism consisting of two similar prisms of calcite. Schematic of the ray path when the incident beam is (a) *z* polarized and (b) *y* polarized. (c) For an unpolarized beam, incident normally, the *z*- and *y*-polarized components will split at the output end.

4.8 Circular Birefringence or Optical Activity

The polarization eigenmodes of a circularly birefringent or optically active substance are *left-circularly polarized* (LCP) and *right-circularly polarized* (RCP) waves, which propagate with slightly different velocities. This means that if an LCP light beam is incident on the substance, it will propagate as an LCP beam; similarly, an RCP light beam will propagate as an RCP beam but with a slightly different velocity. On the other hand, if a linearly polarized light beam is incident, we must express the linear polarization as a superposition of an LCP and an RCP wave and then consider the independent propagation of the two waves. We illustrate through an example.

Let us consider an LCP and an RCP wave, both of the same amplitudes, propagating in the +*z* direction. The two waves can be described by

LCP:
$$E_x^l = E_0 \cos(\omega t - k_l z)$$
$$E_y^l = E_0 \sin(\omega t - k_l z) \tag{4.59}$$

RCP:
$$E_x^r = E_0 \cos(\omega t - k_r z)$$
$$E_y^r = -E_0 \sin(\omega t - k_r z), \tag{4.60}$$

where $k_{l,r} = (\omega/c)\, n_{l,r}$. The superscript (and the subscript) *l* and *r* represent an LCP and an RCP wave, respectively, and n_l and n_r are the corresponding refractive indices.

If we assume the simultaneous propagation of the two beams, the *x* and *y* components of the resultant fields are given by the following equations:

$$E_x = E_0 [\cos(\omega t - k_r z) + \cos(\omega t - k_l z)]$$

or

$$E_x = 2E_0 \cos\left[\frac{1}{2}(k_r - k_l)\, z\right] \cos[\omega t - \delta(z)]. \tag{4.61}$$

Similarly,

$$E_y = 2E_0 \sin\left[\frac{1}{2}(k_r - k_l)\, z\right] \cos[\omega t - \delta(z)], \tag{4.62}$$

where

$$\delta(z) = \frac{1}{2}(k_l + k_r)\, z .$$

Equations (4.61) and (4.62) represent a linearly polarized beam making an angle $\theta(z)$, with the *x* axis given by

$$\theta(z) = \frac{1}{2}(k_r - k_l)\, z = \frac{\pi}{\lambda}(n_r - n_l)\, z\,. \tag{4.63}$$

Thus, the resultant wave is *always* linearly polarized with the plane of polarization rotating with z. It should be noted that if $n_r > n_l$, then $\theta(z)$ increases with z, so the rotation will be counterclockwise (left), whereas for $n_r < n_l$, the rotation will be clockwise (right). We define the rotation as clockwise or counterclockwise as seen by an observer looking into the source. Depending on the relative values of n_l and n_r, a circularly birefringent medium is classified as follows:

Left-handed, or laevorotatory: $n_r > n_l$
Right-handed, or dextrorotatory: $n_r < n_l$

The optical activity in a sugar solution is due to the helical structure of sugar molecules. The method of determining the concentration of sugar solutions by measuring the rotation of the plane of polarization is a widely used method in industry. Optical activity is also exhibited in crystals. For example, for a linearly polarized light propagating along the optic axis of a quartz crystal, the plane of polarization is rotated. (When a wave propagates along the optic axis of a quartz crystal, it is, strictly speaking, not like calcite. The modes are *not* linearly polarized; they are RCP and LCP propagating with slightly different velocities.) Indeed,

$$\left|n_l - n_r\right| \approx 7\times 10^{-5} \Rightarrow \theta \approx \frac{7}{60}\pi \Rightarrow \theta = 21 \text{ deg} \quad \text{for } z = 0.1 \text{ cm};\ \lambda_0 = 6000 \text{ Å}. \tag{4.64}$$

In optical fibers, optical activity is induced if the fiber is twisted; this phenomenon will be discussed in Chapter 8.

4.8.1 The Faraday effect and circular birefringence

Certain dielectrics become circularly birefringent in the presence of magnetic fields. This is known as the *Faraday effect*, which is a magneto-optical phenomenon. The resulting rotation of an incident plane polarized light beam is proportional to the component of the magnetic field in the direction of propagation and is given by

$$\theta = VBd\,, \tag{4.65}$$

where θ (in radians) is the angle of rotation, B (in teslas) is the component of the magnetic induction vector in the direction of propagation, d (in meters) is the distance over which the light and the magnetic field interact, and V (in radians per teslas per meter) is the proportionality constant known as the *Verdet constant*. Terbium gallium garnet (TGG) has extremely high values of V (≈ -40 rad $\mathrm{T^{-1}m^{-1}}$). One of the very important applications of Faraday rotation is in the construction of the device known as the *Faraday isolator*, discussed in the next section.

4.8.2 The Faraday isolator

Faraday isolators (see Fig. 4.10) allow light to pass through in only one direction and are extensively used to avoid optical feedback. In Fig. 4.10, P_1 and P_2 are two linear polarizers with pass axes at 45 deg to each other. The Faraday rotator is chosen to give a 45-deg rotation. The light beam incident from the left is polarized along the *x* direction. The *x*-polarized light passes through the Faraday rotator, which rotates the SOP by 45 deg. Thus, the beam coming out of the

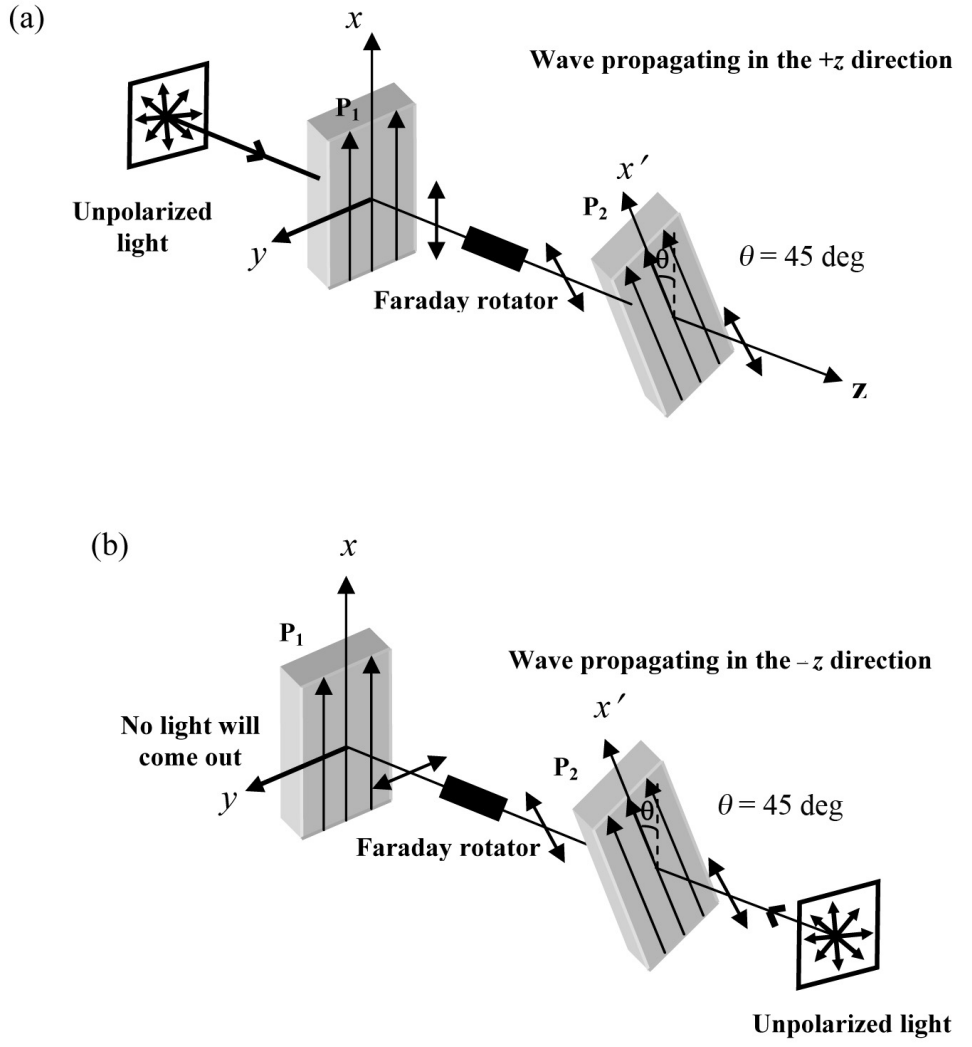


Figure 4.10 Schematic of an optical isolator consisting of two linear polarizers P_1 and P_2, with their pass axes at 45 deg to each other and a Faraday rotator between the polarizers. The Faraday rotator rotates the SOP by 45 deg. (a) For a light beam that is incident from the left, its *x* component will pass through P_2. (b) For a light beam incident from the right, no light will pass through P_1.

Faraday rotator is polarized along the x' direction and is along the pass axis of P_2 [see Fig. 4.10(a)]. The light passes through, and for a good isolator, the transmission can be very high. If a light beam is incident from the right, it will be polarized along the x' direction. The x'-polarized light passes through the Faraday rotator, further rotating the state of polarization by 45 deg. The beam coming out of the Faraday rotator is polarized along the y direction and is perpendicular to the pass axis of P_1; therefore, no light will pass through P_1 [see Fig. 4.10(b)]. We must note that if the magnetic field is along the $+z$ direction, for the wave propagating along the $+z$ direction [see Fig. 4.10(a)], the rotation will be in the counterclockwise direction. On the other hand, for the wave propagating along the $-z$ direction [see Fig. 4.10(a)], the magnetic field is opposite to the direction of propagation, and the Faraday rotation will be in the clockwise direction.

In the wavelength region 0.7–1.1 μm, one often uses terbium-doped borosilicate glass. Faraday isolators are extensively used in many fiber optic devices, and in the wavelength range 1.3–1.55 μm (which is the wavelength range of interest in fiber optic communication systems), one often uses yttrium iron garnet (YIG) crystals.

References

1. A. Ghatak, *Optics*, McGraw-Hill, New York (2009).
2. A. Ghatak and K. Thyagarajan, *Optical Electronics*, Cambridge University Press, New York (1989).

Bibliography

Alonso, M. and E. J. Finn, *Physics*, Addison-Wesley, Reading, MA (1970).

Bird, G. R. and M. P. Parrish, "The wire grid as a near infrared polarizer," *J. Opt. Soc. Am* **50**, 886–891 (1960).

Born, M. and E. Wolf, *Principles of Optics*, Pergamon Press, Oxford, UK (1970).

Chandrashekhar, S., "Simple model for optical activity," *Am. J. Phys.* **24**, 503–506 (1956).

Feynman, R. P., R. B. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Vol. I, Addison-Wesley, Reading, MA (1963).

Gay, P., *An Introduction to Crystal Optics*, Longmans Green and Co., London (1967).

Ghatak, A. and K. Thyagarajan, *Introduction to Fiber Optics*, Cambridge University Press, New York (1998).

Goldstein, D., *Polarized Light*, 2nd ed., Marcel Dekker, New York (2003)

Jenkins, F. A. and H. E. White, *Fundamentals of Optics*, McGraw-Hill, New York (1976).

Shurcliff, W. A. and S. S. Ballard, *Polarized Light*, Van Nostrand, Princeton, NJ (1964).

Waterman, T. H., “Polarized light and animal navigation,” *Sci. Am.* **193**(1) 88–94 (1955).

Wood, E. A., *Crystals and Light*, Van Nostrand Momentum Book No. 5, Van Nostrand, Princeton, NJ (1964).

Chapter 5
Jones Vector Representation of Polarized Light

5.1 Introduction

In Chapter 3, we discussed the basic concepts involved in the polarization of light waves. In various applications of polarized light, one often needs to know how the polarization of the incident wave changes when it is passed through an optical component such as a polarizer or a birefringent medium. This can be described by three formalisms—namely, Jones calculus, Mueller calculus, and the Poincaré sphere. In this chapter, we will discuss the Jones calculus approach, which is simpler than the Mueller calculus. It represents a given SOP by only a two-component vector known as a *Jones vector* and uses simple 2 × 2 matrices to calculate the effect of a polarizer or a birefringent medium on a given SOP.

5.2 Jones Vectors

As discussed in Chapter 3, the electric field vector of an arbitrarily polarized light wave can be described in terms of two orthogonal and linearly polarized components as

$$\mathbf{E}(z,t) = \hat{\mathbf{x}}\, a_x \cos\left(\omega t - kz + \delta_x\right) + \hat{\mathbf{y}}\, a_y \cos\left(\omega t - kz + \delta_y\right). \qquad (5.1)$$

According to the Jones vector representation, the preceding SOP is represented in terms of a 2 × 1 matrix (column vector) as follows:

$$\mathbf{E}(z,t) = \begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} a_x \cos\left(\omega t - kz + \delta_x\right) \\ a_y \cos\left(\omega t - kz + \delta_y\right) \end{bmatrix}. \qquad (5.2)$$

Using the complex variable notations, the preceding can also be written as

$$\mathbf{E}(z,t) = \begin{bmatrix} a_x e^{i(\omega t - kz + \delta_x)} \\ a_y e^{i(\omega t - kz + \delta_y)} \end{bmatrix} = e^{i(\omega t - kz + \delta_x)} \begin{bmatrix} a_x \\ a_y e^{i\delta} \end{bmatrix}, \tag{5.3}$$

where $\delta = (\delta_y - \delta_x)$ represents the relative phase difference between the *x*- and *y*-polarized components. In order to write the preceding equation in a concise form, the common phase factor is omitted, and $\mathbf{E}(z, t)$ is written as

$$\mathbf{E}(z,t) = \begin{bmatrix} a_x \\ a_y e^{i\delta} \end{bmatrix}, \tag{5.4}$$

which is known as the *Jones vector* of the wave, described by Eq. (5.1). The intensity of the wave is given by

$$I = a_x^2 + a_y^2 = E_x^* E_x + E_y^* E_y ,$$

which can be obtained from the Jones vector as follows:

$$I = \mathbf{E}^\dagger \mathbf{E}, \tag{5.5}$$

where $\mathbf{E}^\dagger = \begin{bmatrix} E_x^*, E_y^* \end{bmatrix}$ represents the Hermitian adjoint of $\mathbf{E}$. The Jones vectors $\mathbf{E_1}$ and $\mathbf{E_2}$ are said to be *orthogonal* if they satisfy the condition

$$\mathbf{E_1}^\dagger \mathbf{E_2} = \mathbf{E_2}^\dagger \mathbf{E_1} = 0 . \tag{5.6}$$

5.2.1 Normalized form of Jones vectors

The Jones vector given by Eq. (5.4) is said to be normalized if the amplitudes a_x and a_y are divided by $\sqrt{a_x^2 + a_y^2}$ so that the intensity of the wave is unity. The normalized form of the Jones vector of the wave described by Eq. (5.1) is thus given by

$$\mathbf{E}(z,t) = \frac{1}{\sqrt{a_x^2 + a_y^2}} \begin{bmatrix} a_x \\ a_y e^{i\delta} \end{bmatrix}. \tag{5.7}$$

In the following section, we provide the normalized forms of the Jones vectors that correspond to some common states of polarization.

5.2.1.1 Linear polarization

As discussed in Section 3.2.1, a linearly polarized wave of amplitude a making an angle θ with the x axis can be described as a superposition of the x- and y-polarized components with amplitudes $a\cos\theta$ and $a\sin\theta$, respectively, and $\delta = 0$. Thus, the corresponding normalized Jones vector is given by

$$\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}. \tag{5.8}$$

In the special case of $\theta = 0$ or $\theta = \pi/2$, i.e., a linear horizontal or vertical polarization, the Jones vectors will become

Horizontal: $$\hat{\mathbf{e}}_{\mathbf{x}} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{5.9}$$

Vertical: $$\hat{\mathbf{e}}_{\mathbf{y}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{5.10}$$

Obviously,

$$\hat{\mathbf{e}}_{\mathbf{x}}^{\dagger} \cdot \hat{\mathbf{e}}_{\mathbf{x}} = [1 \quad 0]\begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1,$$

and

$$\hat{\mathbf{e}}_{\mathbf{x}}^{\dagger} \cdot \hat{\mathbf{e}}_{\mathbf{y}} = 0.$$

5.2.1.2 Circular polarization

A left- (right-) circularly polarized wave can be described as a superposition of the x- and y-polarized components having equal amplitudes and the phase difference $\delta = -(+)\ \pi/2$. Thus [using Eq. (5.7)], the Jones vectors of the left- and right-circular basis vectors will be given by

Left-circular: $$\hat{\mathbf{e}}_{\mathbf{LC}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1e^{-\frac{i\pi}{2}} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}, \tag{5.11a}$$

Right-circular: $$\hat{\mathbf{e}}_{\mathbf{RC}} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1e^{+\frac{i\pi}{2}} \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ +i \end{bmatrix}. \tag{5.11b}$$

Obviously,

$$\hat{\mathbf{e}}_{\mathbf{LC}} = \frac{1}{\sqrt{2}}\left(\hat{\mathbf{e}}_{\mathbf{x}} - i\,\hat{\mathbf{e}}_{\mathbf{y}}\right), \tag{5.12a}$$

and

$$\hat{\mathbf{e}}_{\mathbf{RC}} = \frac{1}{\sqrt{2}}\left(\hat{\mathbf{e}}_{\mathbf{x}} + i\,\hat{\mathbf{e}}_{\mathbf{y}}\right). \tag{5.12b}$$

Further,

$$\hat{\mathbf{e}}^{\dagger}_{\mathbf{LC}}.\hat{\mathbf{e}}_{\mathbf{LC}} = 1 = \hat{\mathbf{e}}^{\dagger}_{\mathbf{RC}}.\hat{\mathbf{e}}_{\mathbf{RC}},$$

$$\hat{\mathbf{e}}^{\dagger}_{\mathbf{LC}}.\hat{\mathbf{e}}_{\mathbf{RC}} = \frac{1}{2}\begin{bmatrix}1 & i\end{bmatrix}\begin{bmatrix}1\\ i\end{bmatrix} = 0.$$

5.2.1.3 Elliptical polarization with major and minor axes along the *x* and *y* directions

For a left-elliptically polarized (LEP) wave [see Fig. 5.1(a)] with major and minor axes along the *x* and *y* directions, the electric field can be written as

$$\begin{aligned} E_x &= a\cos\left(\omega t - kz\right), \\ E_y &= b\cos\left(\omega t - kz - \frac{\pi}{2}\right). \end{aligned} \tag{5.13}$$

The normalized Jones vector in terms of its *x* and *y* components will thus be given by

$$\begin{bmatrix}E_x\\ E_y\end{bmatrix} = \frac{1}{\sqrt{a^2+b^2}}\begin{bmatrix}a\\ be^{-\frac{i\pi}{2}}\end{bmatrix} = \begin{bmatrix}\cos\varepsilon\\ -i\sin\varepsilon\end{bmatrix}, \tag{5.14}$$

where $\varepsilon = \tan^{-1}(b/a)$ is a measure of the ellipticity of the wave. Similarly, for a right-elliptically polarized (REP) wave [see Fig. 5.1(b)], the normalized Jones vector will be given by

$$\begin{bmatrix}E_x\\ E_y\end{bmatrix} = \frac{1}{\sqrt{a^2+b^2}}\begin{bmatrix}a\\ be^{+\frac{i\pi}{2}}\end{bmatrix} = \begin{bmatrix}\cos\varepsilon\\ i\sin\varepsilon\end{bmatrix}. \tag{5.15a}$$

It should be noted that the SOP given by Eq. (5.15a) and shown in Fig. 5.1(b) is not orthogonal to the one given by Eq. (5.14) and shown in Fig. 5.1(a). The orthogonal SOP is shown in Fig. 5.1(c), which is right elliptical and whose major axis is along the *y* axis. Mathematically, this SOP is given by

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = \frac{1}{\sqrt{a^2 + b^2}} \begin{bmatrix} b \\ ae^{+\frac{i\pi}{2}} \end{bmatrix} = \begin{bmatrix} \sin\varepsilon \\ i\cos\varepsilon \end{bmatrix}. \tag{5.15b}$$

The orthogonality can be tested using the condition given by Eq. (5.6), as discussed here:

Let $\hat{\mathbf{e}}_1 = \begin{bmatrix} \cos\varepsilon \\ -i\sin\varepsilon \end{bmatrix}$, $\hat{\mathbf{e}}_2 = \begin{bmatrix} \cos\varepsilon \\ i\sin\varepsilon \end{bmatrix}$, and $\hat{\mathbf{e}}_3 = \begin{bmatrix} \sin\varepsilon \\ i\cos\varepsilon \end{bmatrix}$.

Then, $$\hat{\mathbf{e}}_1^{\dagger}\hat{\mathbf{e}}_3 = \begin{bmatrix} \cos\varepsilon & i\sin\varepsilon \end{bmatrix} \begin{bmatrix} \sin\varepsilon \\ i\cos\varepsilon \end{bmatrix} = 0,$$

showing that $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_3$ are orthogonal, while

$$\hat{\mathbf{e}}_1^{\dagger}\hat{\mathbf{e}}_2 = \begin{bmatrix} \cos\varepsilon & i\sin\varepsilon \end{bmatrix} \begin{bmatrix} \cos\varepsilon \\ i\sin\varepsilon \end{bmatrix} = \cos 2\varepsilon,$$

showing that $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ are not orthogonal, except when $\varepsilon = \pi/4$, in which case, $\hat{\mathbf{e}}_2$ and $\hat{\mathbf{e}}_3$ both represent the same SOP—namely, right-circularly polarized.

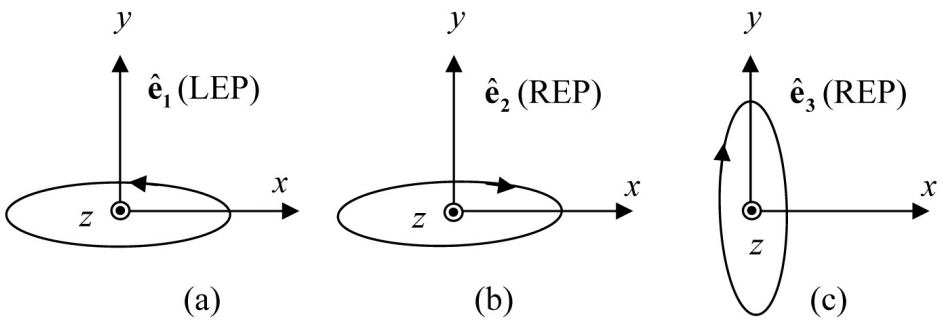


Figure 5.1 A typical (a) left- and (b) right-elliptically polarized beam with its major axis along the *x* axis. (c) Right-elliptical SOP that is orthogonal to the one shown in (a). Propagation is along the *z* direction, which is coming out of the page.

5.2.1.4 General elliptical polarization

Let us consider a general elliptically polarized wave [shown in Figs. 5.2(a) and (b)] with its major and minor axes along the ξ and η directions such that the angle between the ξ and x axes is θ. As discussed earlier, this wave can be described in terms of ξ and η components as

$$\begin{aligned} E_\xi &= a\cos\left(\omega t - kz\right) \\ E_\eta &= b\cos\left(\omega t - kz \mp \frac{\pi}{2}\right), \end{aligned} \tag{5.16}$$

where the upper and lower sign corresponds to the left and right sense of rotation of the electric field vector. The normalized Jones vector in terms of its ξ and η components will thus be given by

$$\begin{bmatrix} E_\xi \\ E_\eta \end{bmatrix} = \frac{1}{\sqrt{a^2+b^2}} \begin{bmatrix} a \\ be^{\mp\frac{i\pi}{2}} \end{bmatrix} = \begin{bmatrix} \cos\varepsilon \\ \mp i\sin\varepsilon \end{bmatrix}. \tag{5.17}$$

The Jones vector in terms of the x- and y-polarized components can now be obtained by rotating the coordinate system clockwise by an angle θ:

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\varepsilon \\ \mp i\sin\varepsilon \end{bmatrix} = \begin{bmatrix} \cos\theta\cos\varepsilon \pm i\sin\theta\sin\varepsilon \\ \sin\theta\cos\varepsilon \mp i\cos\theta\sin\varepsilon \end{bmatrix}. \tag{5.18}$$

Thus, the normalized Jones vectors for a left- and right-elliptically polarized wave whose representative ellipses make an angle θ with the x axis and have major and minor axes as $2a$ and $2b$ are given by

$$\hat{\mathbf{e}}_{\mathbf{LE}} = \begin{bmatrix} \cos\theta\cos\varepsilon + i\sin\theta\sin\varepsilon \\ \sin\theta\cos\varepsilon - i\cos\theta\sin\varepsilon \end{bmatrix}, \tag{5.19}$$

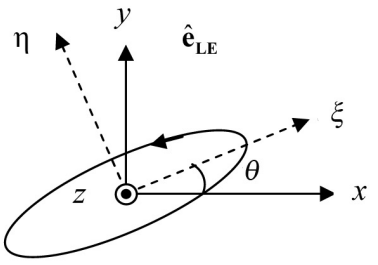
and

$$\hat{\mathbf{e}}'_{\mathbf{RE}} = \begin{bmatrix} \cos\theta\cos\varepsilon - i\sin\theta\sin\varepsilon \\ \sin\theta\cos\varepsilon + i\cos\theta\sin\varepsilon \end{bmatrix}, \tag{5.20}$$

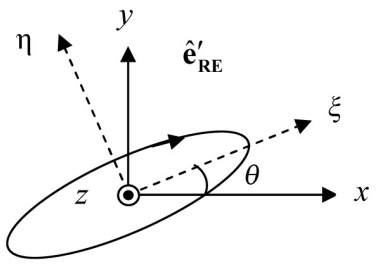
respectively. It should be noted again that the SOP given by Eq. (5.20) and shown in Fig. 5.2(b) is not orthogonal to the one given by Eq. (5.19) and shown in Fig. 5.2(a). This can be easily checked for the θ = 0 case, for which the preceding states of polarization are given by

$$\begin{bmatrix} \cos\varepsilon \\ -i\sin\varepsilon \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \cos\varepsilon \\ +i\sin\varepsilon \end{bmatrix}, \tag{5.21}$$

which are not orthogonal, as discussed earlier.

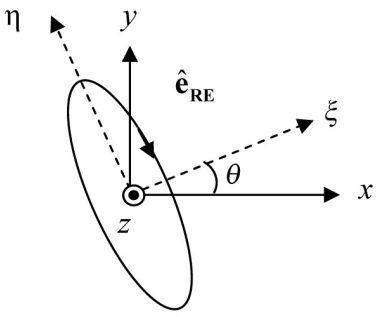


(a)



$$\hat{\mathbf{e}}'^{\dagger}_{\mathbf{RE}} \cdot \hat{\mathbf{e}}_{\mathbf{LE}} \neq 0$$

(b)



$$\hat{\mathbf{e}}^{\dagger}_{\mathbf{RE}} \cdot \hat{\mathbf{e}}_{\mathbf{LE}} = 0$$

(c)

Figure 5.2 Typical (a) left- and (b)**,** (c) right-elliptically polarized beams described by Eqs. (5.19), (5.20), and (5.22), respectively. The beam shown in (c) is orthogonal to that shown in (a). Propagation is along the *z* direction, which is coming out of the page.

The SOP that is orthogonal to the one shown in Fig. 5.2(a) will have its minor axis oriented at θ, as shown in Fig. 5.2(c); the Jones vector of this SOP can be obtained from Eq. (5.20) by interchanging the terms cos ε and sin ε, and is given by

$$\hat{\mathbf{e}}_{\mathbf{RE}} = \begin{bmatrix} \cos\theta \sin\varepsilon - i \sin\theta \cos\varepsilon \\ \sin\theta \sin\varepsilon + i \cos\theta \cos\varepsilon \end{bmatrix}. \tag{5.22}$$

Thus, the two orthogonal and elliptically polarized basis vectors are given by Eqs. (5.19) and (5.22), respectively. We will use these again when deriving the Jones vector of a given SOP in terms of elliptically polarized basis vectors.

5.3 Jones Matrices

The beauty of the Jones vector representation is that the effect of any polarization component on a given SOP can be obtained by multiplying the input Jones vector by a 2 × 2 matrix known as the *Jones matrix* of the component. Thus, if a given polarization state **A** is passed through a polarizing device whose Jones matrix is J, the output polarization state $\mathbf{A}'$ will be given by $\mathbf{A}' = J\,\mathbf{A}$.

If there are two devices in a series with Jones matrices J_1 and J_2, respectively, the Jones matrix of the combination is given by $J_2\,J_1$, and the output polarization state will be given by $\mathbf{A}' = J_2\,J_1\,\mathbf{A}$. In the following section, we obtain Jones matrices for some basic polarization components.

5.3.1 Linear polarizer

A *linear polarizer* is a device that allows only one of the two orthogonal linearly polarized components of a given SOP to pass through it. The direction of polarization of the passed component is known as the *pass axis* of the polarizer. Let us consider a linear polarizer whose pass axis makes an angle θ with the x axis (Fig. 5.3).

If a light beam with x and y components E_x and E_y is passed through the given polarizer, the output SOP will be linearly polarized with its electric field given by

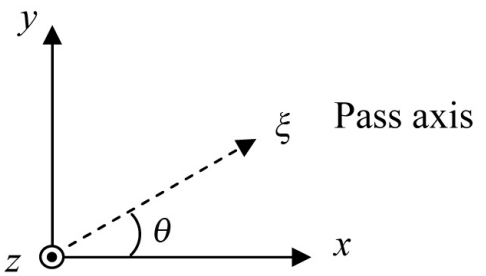


Figure 5.3 A linear polarizer with pass axis at an angle θ to the x axis.

$$\mathbf{E}' = \hat{\xi}(E_x \cos\theta + E_y \sin\theta), \tag{5.23}$$

where $\hat{\xi}$ represents a unit vector along the pass axis ξ. The x and y components of the output SOP are thus given by

$$\begin{aligned} E'_x &= (E_x \cos\theta + E_y \sin\theta)\cos\theta, \\ \text{and} \quad E'_y &= (E_x \cos\theta + E_y \sin\theta)\sin\theta. \end{aligned} \tag{5.24}$$

It is evident from the preceding equations that the Jones vectors of the input and output states of polarization are related as

$$\begin{bmatrix} E'_x \\ E'_y \end{bmatrix} = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix} \begin{bmatrix} E_x \\ E_y \end{bmatrix}. \tag{5.25}$$

Thus, the Jones matrix of the given linear polarizer is given by

$$T_{\mathrm{LP}}(\theta) = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}. \tag{5.26}$$

Substituting θ = 0, π/4, or π/2 in Eq. (5.26), one obtains Jones matrices for a linear polarizer with its pass axis oriented along the x, θ = π/4, and y directions, respectively:

$$T_{\mathrm{LP}}(0) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ T_{\mathrm{LP}}(\pi/4) = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \text{ and } T_{\mathrm{LP}}(\pi/2) = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{5.27}$$

5.3.2 Linear retarder

A *linear retarder* splits the input SOP into two orthogonal but linear (slow and fast) components and introduces a certain phase difference δ (retardation) between them. Let the fast axis of the given retarder make an angle θ with the x axis (Fig. 5.4).

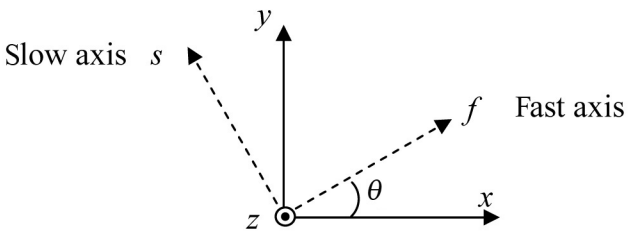


Figure 5.4 A linear retarder with its fast axis at an angle θ to the x axis.

The output SOP corresponding to an incident light beam (E_x, E_y) can be obtained as follows: We first obtain the components of the incident SOP along the fast and slow (*f* and *s*) axes (by rotating the *x*-*y* coordinate system counterclockwise by an angle θ), which are given by

$$\begin{bmatrix} E_{\mathrm{f}} \\ E_{\mathrm{s}} \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} E_x \\ E_y \end{bmatrix}. \tag{5.28}$$

As these components pass through the retarder, the fast component is advanced in phase by δ, and the corresponding components at the output of the retarder are then given by

$$E'_{\mathrm{f}} = e^{i\delta} E_{\mathrm{f}} ,$$

and

$$E'_{\mathrm{s}} = E_s ,$$

i.e.,

$$\begin{bmatrix} E'_{\mathrm{f}} \\ E'_{\mathrm{s}} \end{bmatrix} = \begin{bmatrix} e^{i\delta} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} E_{\mathrm{f}} \\ E_{\mathrm{s}} \end{bmatrix}. \tag{5.29}$$

Once we know E'_f and E'_s, the components along the *x* and *y* axes at the output are obtained by rotating the coordinate axes clockwise by an angle θ. These components are given by

$$\begin{bmatrix} E'_x \\ E'_y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} E'_{\mathrm{f}} \\ E'_{\mathrm{s}} \end{bmatrix}. \tag{5.30}$$

Using Eqs. (5.29) and (5.28) in Eq. (5.30), it can easily be shown that the Jones vectors of the input and the output beams are related through the relation

$$\begin{bmatrix} E'_x \\ E'_y \end{bmatrix} = T_{\mathrm{LR}}(\theta) \begin{bmatrix} E_x \\ E_y \end{bmatrix}, \tag{5.31}$$

where

$$\begin{aligned} T_{\mathrm{LR}}(\theta) &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e^{i\delta} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \\ &= \begin{bmatrix} e^{i\delta}\cos^2\theta + \sin^2\theta & (e^{i\delta}-1)\sin\theta\cos\theta \\ (e^{i\delta}-1)\sin\theta\cos\theta & e^{i\delta}\sin^2\theta + \cos^2\theta \end{bmatrix} \end{aligned} \tag{5.32}$$

represents the Jones matrix of a linear retarder whose fast axis makes an angle θ with the x axis. The most commonly used linear retarders are the half-wave plate (HWP) and the quarter-wave plate (QWP), whose Jones matrices can be obtained from the preceding equation by substituting $\delta = \pi$ and $\pi/2$, respectively. Thus, the Jones matrices for an HWP or a QWP whose fast axis is oriented at angle θ with the x axis are given as

HWP:
$$T_{\text{HWP}}(\theta) = \begin{bmatrix} -\cos 2\theta & -\sin 2\theta \\ -\sin 2\theta & \cos 2\theta \end{bmatrix}, \tag{5.33}$$

QWP:
$$T_{\text{QWP}}(\theta) = \begin{bmatrix} i\cos^2\theta + \sin^2\theta & (i-1)\sin\theta\cos\theta \\ (i-1)\sin\theta\cos\theta & i\sin^2\theta + \cos^2\theta \end{bmatrix}. \tag{5.34}$$

Further, by substituting $\theta = 0$ and $\pi/2$ in Eq. (5.32), one obtains the following Jones matrices for linear retarders, with the fast axis parallel to the x axis and y axis:

$$T_{\text{LR}}(0) = \begin{bmatrix} e^{i\delta} & 0 \\ 0 & 1 \end{bmatrix}, \text{ and } T_{\text{LR}}(\pi/2) = \begin{bmatrix} 1 & 0 \\ 0 & e^{i\delta} \end{bmatrix}. \tag{5.35a}$$

Substituting $\delta = \pi/2$ and π in the preceding equation, one obtains the corresponding Jones matrices for a QWP and an HWP, as given by

$$T_{\text{QWP}}(0) = \begin{bmatrix} i & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad T_{\text{QWP}}(\pi/2) = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}, \tag{5.35b}$$

$$T_{\text{HWP}}(0) = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad T_{\text{HWP}}(\pi/2) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \tag{5.35c}$$

Example 5.1 Obtain the output SOP if left-circularly polarized light is passed through a linear polarizer with its pass axis making an angle θ with the x axis.

Solution: The input state is given by

$$\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}.$$

The Jones matrix of the linear polarizer is given by Eq. (5.26) as

$$T_{\mathrm{LP}}(\theta) = \begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix}.$$

The Jones vector of the output SOP will then be

$$\begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \cos\theta\sin\theta & \sin^2\theta \end{bmatrix} \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \cos^2\theta - i\sin\theta\cos\theta \\ \cos\theta\sin\theta - i\sin^2\theta \end{bmatrix}$$

$$= \frac{e^{-i\theta}}{\sqrt{2}} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.$$

As expected, this represents linearly polarized light oriented at θ with half the input intensity.

Example 5.2 Obtain the output SOP when linearly polarized light oriented at θ is passed through a QWP whose fast axis is oriented along the x axis.

Solution: The input state is given by

$$\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix}.$$

The Jones matrix of the given QWP is given by Eq. (5.35b) as

$$T_{\mathrm{QWP}}(0) = \begin{bmatrix} e^{i\pi/2} & 0 \\ 0 & 1 \end{bmatrix}.$$

The Jones vector of the output SOP will then be given by

$$\begin{bmatrix} e^{i\pi/2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \begin{bmatrix} e^{i\pi/2}\cos\theta \\ \sin\theta \end{bmatrix} = e^{i\pi/2} \begin{bmatrix} \cos\theta \\ -i\sin\theta \end{bmatrix}.$$

This represents left-elliptically polarized light whose ellipticity is given by $b/a = \tan\theta$ and whose major axis is oriented along the x axis. If $\theta = \pi/4$, then $b/a = 1$, and hence, the output SOP will be left circular (see Fig. 5.5).

Example 5.3 Obtain the output SOP when linearly polarized light oriented at θ is passed through an HWP whose fast axis is oriented along the x axis.

Solution: The Jones vector of the wave coming out of the HWP will be given by

$$\begin{bmatrix} e^{i\pi} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \begin{bmatrix} -\cos\theta \\ \sin\theta \end{bmatrix} = e^{i\pi} \begin{bmatrix} \cos\theta \\ -\sin\theta \end{bmatrix},$$

which represents a linearly polarized wave making an angle $-\theta$ with the *x* axis (see Fig. 5.6).

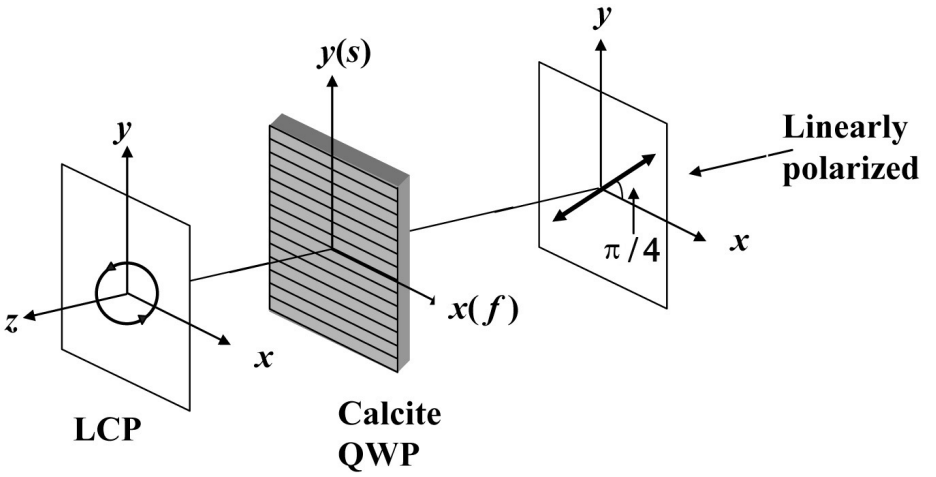


Figure 5.5 If a linearly polarized wave making an angle π /4 with the *x* axis is incident on a QWP (whose fast axis is along the *x* axis), the output beam will be left-circularly polarized.

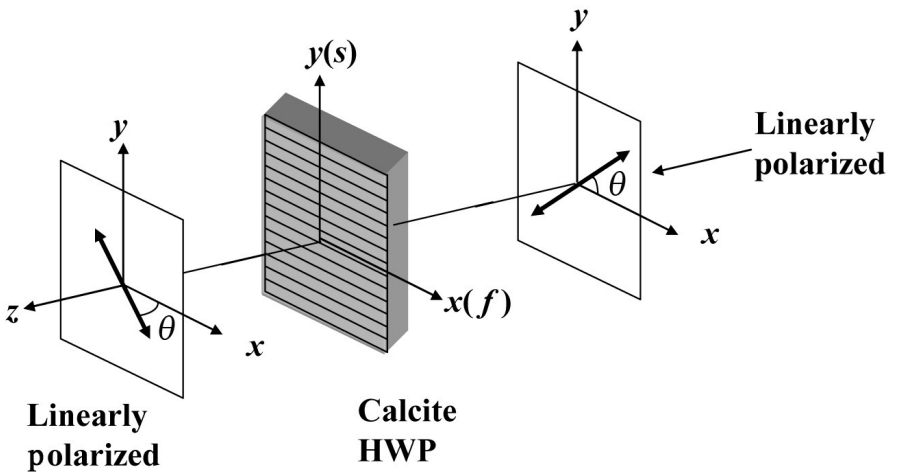


Figure 5.6 If a linearly polarized wave making an angle θ with the *x* axis is incident on an HWP (whose fast axis is along the *x* axis), the output beam will be linearly polarized at an angle $-\theta$.

5.4 Jones Vectors in Terms of the Circular Basis Vectors

In the previous section, we discussed the Jones vector of a given SOP in terms of two orthogonal and linear components, which is relatively simple and is used most of the time. However, if a polarized beam is passed through a circularly birefringent medium (such as a Faraday rotator), where the two eigen-polarization states are left- and right-circularly polarized, it is much more convenient to use Jones vectors in terms of circularly polarized basis vectors. In this section, we discuss how to obtain the left- and right-circularly polarized components (E_{LC}, E_{RC}) of a given SOP, if its x and y components (E_x, E_y) are known.

Since any SOP can be expressed as a superposition of any two orthogonally polarized states, the electric field **E** of a given SOP can also be expressed in terms of its circularly polarized components. Thus,

$$\mathbf{E} = E_x\hat{\mathbf{e}}_{\mathbf{x}} + E_y\hat{\mathbf{e}}_{\mathbf{y}} = E_{\mathrm{LC}}\hat{\mathbf{e}}_{\mathbf{LC}} + E_{\mathrm{RC}}\hat{\mathbf{e}}_{\mathbf{RC}}, \tag{5.36}$$

where $\hat{\mathbf{e}}_{\mathbf{x}}$ and $\hat{\mathbf{e}}_{\mathbf{y}}$ are the linear basis vectors [Eqs. (5.9) and (5.10)], and $\hat{\mathbf{e}}_{\mathbf{LC}}$ and $\hat{\mathbf{e}}_{\mathbf{RC}}$ represent the circular (left and right) basis vectors [Eqs. (5.11a) and (5.11b)]. Equation (5.36) can be written in matrix form as

$$E_x\begin{bmatrix}1\\0\end{bmatrix} + E_y\begin{bmatrix}0\\1\end{bmatrix} = \frac{E_{\mathrm{LC}}}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix} + \frac{E_{\mathrm{RC}}}{\sqrt{2}}\begin{bmatrix}1\\i\end{bmatrix},$$

or

$$\begin{bmatrix}E_x\\E_y\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\-i & i\end{bmatrix}\begin{bmatrix}E_{\mathrm{LC}}\\E_{\mathrm{RC}}\end{bmatrix} = S\begin{bmatrix}E_{\mathrm{LC}}\\E_{\mathrm{RC}}\end{bmatrix}, \tag{5.37}$$

where

$$S = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\-i & i\end{bmatrix}. \tag{5.38}$$

Multiplying Eq. (5.37) by S^{-1} (the inverse of S), one obtains

$$\begin{bmatrix}E_{\mathrm{LC}}\\E_{\mathrm{RC}}\end{bmatrix} = S^{-1}\begin{bmatrix}E_x\\E_y\end{bmatrix}, \tag{5.39}$$

where

$$S^{-1} = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & i\\1 & -i\end{bmatrix}. \tag{5.40}$$

Thus, by using matrices S^{-1} and S, one can obtain the circularly polarized components from the linearly polarized components of a given SOP [using Eq. (5.39)] and vice versa [using Eq. (5.37)].

Example 5.4 Consider an arbitrary SOP described by the following normalized Jones vector in a linear basis:

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \tag{5.41}$$

where α and β are complex numbers with $|\alpha|^2 + |\beta|^2 = 1$. Obtain its left- and right-circularly polarized components.

Solution: In this example, we will express the preceding SOP in terms of basis vectors $\hat{\mathbf{e}}_{\mathbf{LC}}$ and $\hat{\mathbf{e}}_{\mathbf{RC}}$:

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = c_1 \hat{\mathbf{e}}_{\mathbf{LC}} + c_2 \hat{\mathbf{e}}_{\mathbf{RC}} = c_1 \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} + c_2 \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ +i \end{bmatrix}.$$

Thus, $\quad c_1 + c_2 = \sqrt{2}\,\alpha, \quad$ and $\quad c_1 - c_2 = i\sqrt{2}\,\beta.$

Simple algebra gives

$$c_1 = \frac{1}{\sqrt{2}}(\alpha + i\beta), \text{ and } c_2 = \frac{1}{\sqrt{2}}(\alpha - i\beta).$$

Therefore, c_1 and c_2 are the amplitudes of the left- and right-circular components of the given SOP, and $|c_1|^2$ and $|c_2|^2$ can be interpreted as the fractional powers associated with the LCP and RCP, respectively. Here, c_1 and c_2 are simply the E_{LC} and E_{RC} used in Eq. (5.36). E_{LC} and E_{RC} can also be obtained by multiplying the given Jones vector by S^{-1} [Eq. (5.39)], as shown here:

$$\begin{bmatrix} E_{\text{LC}} \\ E_{\text{RC}} \end{bmatrix} = S^{-1} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \alpha + i\beta \\ \alpha - i\beta \end{bmatrix}.$$

This also shows that the amplitudes of the left- and right-circular components of the given SOP are $\frac{1}{\sqrt{2}}(\alpha + i\beta)$ and $\frac{1}{\sqrt{2}}(\alpha - i\beta)$, respectively.

5.4.1 Jones matrix of an ideal circular polarizer

Let us first consider an ideal left-circular polarizer that would eliminate the right-circularly polarized component of any input SOP. Let the input SOP be given by

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix}.$$

As discussed earlier, the corresponding circularly polarized components (at the input of the polarizer) will be given by Eq. (5.39). The polarizer will eliminate the right-circularly polarized component of the incident SOP; hence, the circularly polarized components of the output SOP will be given by

$$\begin{bmatrix} E'_{\mathrm{LC}} \\ E'_{\mathrm{RC}} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} E_{\mathrm{LC}} \\ E_{\mathrm{RC}} \end{bmatrix}. \tag{5.42}$$

The linearly polarized components of the output SOP can now be obtained using Eq. (5.37) as

$$\begin{bmatrix} E'_x \\ E'_y \end{bmatrix} = S \begin{bmatrix} E'_{\mathrm{LC}} \\ E'_{\mathrm{RC}} \end{bmatrix}. \tag{5.43}$$

Using Eqs. (5.42) and (5.39) in the preceding equation, one obtains

$$\begin{bmatrix} E'_x \\ E'_y \end{bmatrix} = S \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} S^{-1} \begin{bmatrix} E_x \\ E_y \end{bmatrix}. \tag{5.44}$$

Thus, the required Jones matrix of an ideal left-circular polarizer is

$$T_{\mathrm{LCP}} = S \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} S^{-1}. \tag{5.45}$$

Substituting the values of S and S^{-1}, from Eqs. (5.38) and (5.40), respectively, in the preceding equation, one obtains

$$T_{\mathrm{LCP}} = \frac{1}{2} \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}. \tag{5.46}$$

In a similar manner, the Jones matrix of an ideal right-circular polarizer can be derived and given as

$$T_{\text{RCP}} = S \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} S^{-1} = \frac{1}{2} \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}. \tag{5.47}$$

Example 5.5 Show that when a general SOP is passed through an ideal right-circular polarizer, only its right-circular component is passed through.

Solution: Let the incident SOP be given by Eq. (5.41). If this SOP is passed through an ideal right-circular polarizer described by Eq. (5.47), the output SOP will be given by

$$\mathbf{A}' = T_{\text{RCP}} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \alpha - i\beta \\ i\alpha + \beta \end{bmatrix}$$

$$= \frac{\alpha - i\beta}{\sqrt{2}} \; \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} = \frac{\alpha - i\beta}{\sqrt{2}} \; \hat{\mathbf{e}}_{\mathbf{RC}}.$$

Thus, the output SOP is right-circularly polarized with amplitude $(\alpha - i\beta)/\sqrt{2}$, which is simply the right-circular component of the input SOP (see Example 5.4).

Example 5.6 Obtain the Jones matrix of the combination of a QWP with its fast axis along the *x* axis, followed by a linear polarizer (making an angle π/4 with the *x* axis), followed by another QWP with its fast axis parallel to the *y* axis, as shown in Fig. 5.7.

Solution: Using Eqs. (5.35b) and (5.27), the Jones matrix for the previously mentioned QWP + linear polarizer + QWP will be given by

$$T_{\text{QWP}}(\pi/2) T_{\text{LP}}(\pi/4) T_{\text{QWP}}(0) = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} i & 0 \\ 0 & 1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} i & 1 \\ i & 1 \end{bmatrix} = \frac{i}{2} \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}$$

$$= \frac{e^{\frac{i\pi}{2}}}{2} \begin{bmatrix} 1 & -i \\ i & 1 \end{bmatrix}.$$

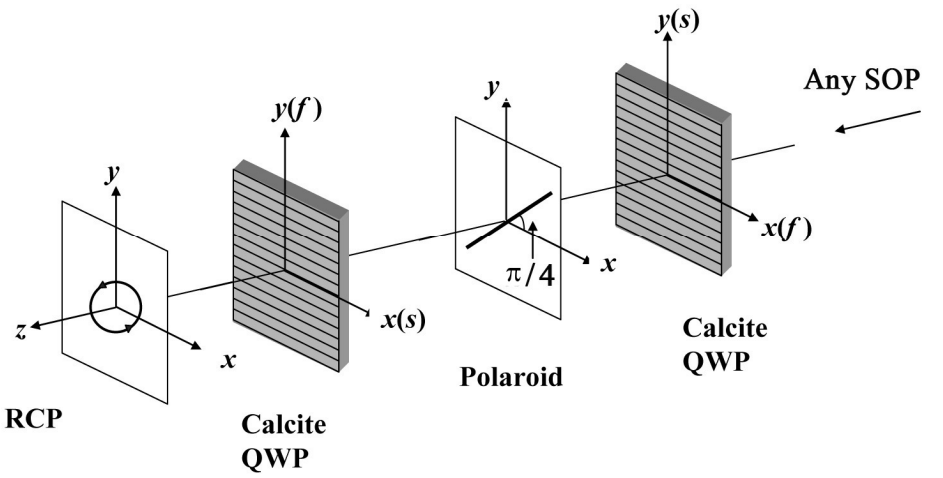


Figure 5.7 A right-circular polarizer will consist of a QWP with its fast axis along the *x* axis, followed by a linear polarizer with its pass axis at an angle π/4 with the *x* axis, followed by another QWP with its fast axis along the *y* axis.

The preceding matrix is the same as that of an ideal right-circular polarizer [see Eq. (5.47)] except for a constant phase factor $e^{i\pi/2}$. Thus, a combination of a QWP with its fast axis along the *x* axis, followed by a linear polarizer (at an angle π/4 with the *x* axis), followed by another QWP with its fast axis parallel to the *y* axis (Fig. 5.7) acts as an ideal right-circular polarizer. Similarly, it can be shown that a combination of a QWP with its fast axis along the *x* axis, followed by a linear polarizer (at an angle −π/4 with the *x* axis), followed by another QWP with its fast axis parallel to the *y* axis acts as an ideal left-circular polarizer.

5.4.2 Jones matrix of a circularly polarized medium (Faraday rotator)

We first consider a circularly birefringent medium in which the left-circular component is the fast component. Let the phase difference introduced between the left- and right-circular components be δ. In order to obtain the Jones matrix of such a medium, we first obtain the left- and right-circular components of the given SOP, introduce the phase difference δ between the two components, and then convert these back to *x* and *y* components. By doing so, it can easily be shown that the output and input states of polarization will be related through the following equation:

$$\begin{bmatrix} E_x' \\ E_y' \end{bmatrix} = S \begin{bmatrix} e^{i\delta} & 0 \\ 0 & 1 \end{bmatrix} S^{-1} \begin{bmatrix} E_x \\ E_y \end{bmatrix}. \tag{5.48}$$

Substituting the values of S and S^{-1} from Eqs. (5.38) and (5.40), respectively, the required Jones matrix is given by

$$S\begin{bmatrix} e^{i\delta} & 0 \\ 0 & 1 \end{bmatrix} S^{-1} = e^{i\delta/2}\begin{bmatrix} \cos\delta/2 & -\sin\delta/2 \\ \sin\delta/2 & \cos\delta/2 \end{bmatrix}. \tag{5.49}$$

Ignoring the constant phase factor $e^{i\delta/2}$, the Jones matrix of a circularly birefringent medium with the left component as the fast component (left-circular retarder) is given by

$$T_{\text{LCR}} = \begin{bmatrix} \cos\delta/2 & -\sin\delta/2 \\ \sin\delta/2 & \cos\delta/2 \end{bmatrix}. \tag{5.50}$$

The Jones matrix of a right-circular retarder can be obtained by changing the sign of δ from positive to negative, which is given as

$$T_{\text{RCR}} = \begin{bmatrix} \cos\delta/2 & \sin\delta/2 \\ -\sin\delta/2 & \cos\delta/2 \end{bmatrix}. \tag{5.51}$$

Example 5.7 Obtain the output SOP when linearly polarized light oriented at θ is passed through a circularly birefringent medium, introducing a phase difference δ between the two components.

Solution: The input SOP will be represented by Eq. (5.8), and, assuming that the left-circular component is the fast component, the Jones matrix of the medium is given by Eq. (5.50). The output SOP will thus be given by

$$\begin{bmatrix} \cos\delta/2 & -\sin\delta/2 \\ \sin\delta/2 & \cos\delta/2 \end{bmatrix}\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} = \begin{bmatrix} \cos(\theta+\delta/2) \\ \sin(\theta+\delta/2) \end{bmatrix},$$

which represents a linearly polarized wave rotated by an angle δ /2 with respect to the input SOP.

Example 5.8 Obtain the output SOP when elliptically polarized light oriented at θ is passed through a circularly birefringent medium, introducing a phase difference δ between the two components.

Solution: Let the input SOP be represented by Eq. (5.18). Assuming that the left-circular component is the fast component, the Jones matrix of the medium is given by Eq. (5.50). The output SOP will thus be given by

$$\begin{bmatrix} \cos\delta/2 & -\sin\delta/2 \\ \sin\delta/2 & \cos\delta/2 \end{bmatrix}\begin{bmatrix} \cos\theta\cos\varepsilon \pm i\sin\theta\sin\varepsilon \\ \sin\theta\cos\varepsilon \mp i\cos\theta\sin\varepsilon \end{bmatrix}$$

$$= \begin{bmatrix} \cos(\theta+\delta/2)\cos\varepsilon \pm i\sin(\theta+\delta/2)\sin\varepsilon \\ \sin(\theta+\delta/2)\cos\varepsilon \mp i\cos(\theta+\delta/2)\sin\varepsilon \end{bmatrix}.$$

The preceding Jones vector represents an elliptically polarized wave having the same ellipticity and sense of rotation as that of the input SOP, except that its major axis is rotated by an angle δ /2. Such a medium thus acts as a *rotator*.

Example 5.9 Show that a left-circularly polarized beam will pass through a left-circular polarizer completely and will be stopped if passed through a right-circular polarizer.

Solution: The Jones vector of a left-circularly polarized beam of unit intensity is given by $\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix}$, and the Jones matrices of the left- and right-circular polarizers are given by

$$T_{\text{LCP}} = \frac{1}{2}\begin{bmatrix}1 & i\\-i & 1\end{bmatrix}, \quad \text{and} \quad T_{\text{RCP}} = \frac{1}{2}\begin{bmatrix}1 & -i\\i & 1\end{bmatrix},$$

respectively. Thus, the output SOP corresponding to the left-circular polarizer will be

$$\frac{1}{2}\begin{bmatrix}1 & i\\-i & 1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix} = \frac{1}{2\sqrt{2}}\begin{bmatrix}2\\-2i\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix},$$

which is the input SOP itself of unit intensity. On the other hand, the output SOP corresponding to the right-circular polarizer will be

$$\frac{1}{2}\begin{bmatrix}1 & -i\\i & 1\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-i\end{bmatrix} = \frac{1}{2\sqrt{2}}\begin{bmatrix}0\\0\end{bmatrix},$$

representing no output.

5.5 Jones Vectors in Terms of the Elliptical Basis Vectors

In general, a medium may be elliptically birefringent, e.g., an optical fiber exposed to both linear and circular birefringence. The two eigenpolarization states of the medium will then be elliptically polarized; in this case, it is much more convenient to use Jones vectors expressed in terms of elliptical basis vectors. This can be done in a similar way as we did earlier for circular basis vectors.

Let E_{LE} and E_{RE} represent the left- and the right-elliptical components of a given SOP whose x and y components are E_x and E_y, respectively. The total field can then be expressed as

$$\mathbf{E} = E_x\hat{\mathbf{e}}_\mathbf{x} + E_y\hat{\mathbf{e}}_\mathbf{y} = E_{\mathrm{LE}}\,\hat{\mathbf{e}}_\mathbf{LE} + E_{\mathrm{RE}}\,\hat{\mathbf{e}}_\mathbf{RE}, \tag{5.52}$$

where $\hat{e}_{\mathrm{LE}}$ and $\hat{e}_{\mathrm{RE}}$ represent the left- and right-elliptical basis vectors given by Eqs. (5.19) and (5.22), respectively. Equation (5.51) can be written in matrix form as

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix} = E_{\mathrm{LE}}\begin{bmatrix} \cos\theta\cos\varepsilon + i\sin\theta\sin\varepsilon \\ \sin\theta\cos\varepsilon - i\cos\theta\sin\varepsilon \end{bmatrix} + E_{\mathrm{RE}}\begin{bmatrix} \cos\theta\sin\varepsilon - i\sin\theta\cos\varepsilon \\ \sin\theta\sin\varepsilon + i\cos\theta\cos\varepsilon \end{bmatrix}$$

$$= R\begin{bmatrix} E_{\mathrm{LE}} \\ E_{\mathrm{RE}} \end{bmatrix}, \tag{5.53}$$

where

$$R = \begin{bmatrix} \cos\theta\cos\varepsilon + i\sin\theta\sin\varepsilon & \cos\theta\sin\varepsilon - i\sin\theta\cos\varepsilon \\ \sin\theta\cos\varepsilon - i\cos\theta\sin\varepsilon & \sin\theta\sin\varepsilon + i\cos\theta\cos\varepsilon \end{bmatrix}. \tag{5.54}$$

Multiplying Eq. (5.53) by R^{-1}, one obtains

$$\begin{bmatrix} E_{\mathrm{LE}} \\ E_{\mathrm{RE}} \end{bmatrix} = R^{-1}\begin{bmatrix} E_x \\ E_y \end{bmatrix}, \tag{5.55}$$

where

$$R^{-1} = \begin{bmatrix} \cos\theta\cos\varepsilon - i\sin\theta\sin\varepsilon & \sin\theta\cos\varepsilon + i\cos\theta\sin\varepsilon \\ \cos\theta\sin\varepsilon + i\sin\theta\cos\varepsilon & \sin\theta\sin\varepsilon - i\cos\theta\cos\varepsilon \end{bmatrix}. \tag{5.56}$$

Thus, by using matrices R^{-1} and R, one can obtain the elliptically polarized components from the linearly polarized components of a given SOP and vice versa.

5.5.1 Jones matrix of an ideal elliptical polarizer and retarder

Jones matrices for an ideal elliptical polarizer and retarder can be obtained in a similar way, using matrices R^{-1} and R [given by Eqs. (5.56) and (5.54)], as we obtained for an ideal circular polarizer and retarder using matrices S^{-1} and S. These matrices are given by the following expressions:

Left- and right-elliptical polarizers:

$$T_{\mathrm{LEP}} = R \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} R^{-1} \tag{5.57}$$

$$T_{\mathrm{REP}} = R \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} R^{-1} \tag{5.58}$$

Elliptical retarder with its left elliptic component as the fast component:

$$T_{\mathrm{LER}} = R \begin{bmatrix} e^{i\delta} & 0 \\ 0 & 0 \end{bmatrix} R^{-1} \tag{5.59}$$

Elliptical retarder with its right elliptic component as the fast component:

$$T_{\mathrm{RER}} = R \begin{bmatrix} 0 & 0 \\ 0 & e^{i\delta} \end{bmatrix} R^{-1} \tag{5.60}$$

Bibliography

Azzam, R. M. A. and N. M. Bashara, *Ellipsometry and Polarized Light*, North Holland Publishing Company, Amsterdam (1977).

Collett, E., *Field Guide to Polarization*, SPIE Press, Bellingham, WA (2005). [doi: 10.1117/3.626141].

Jones, R. C., "New calculus for the treatment of optical systems I. Description and discussion of the calculus," *J. Opt. Soc. Am.* **31**, 488–493 (1941).

Shurcliff, W. A., *Polarized Light: Production and Use*, Harvard University Press, Cambridge, MA (1962).

Theocaris, P. S. and E. E. Gdoutos, *Matrix Theory of Photoelasticity*, Springer Verlag, Berlin–Heidelberg (1979).

Chapter 6
The Stokes Parameters Representation

6.1 Introduction

In Chapter 5, we examined the Jones vector representation of polarized light, which is quite simple and straightforward. This representation, however, can describe only polarized light. In this chapter, we discuss a more versatile representation of the polarization state of any light wave (completely polarized, unpolarized, or partially polarized), which was first proposed by George Gabriel Stokes in 1852.

6.2 The Stokes Parameters

The representation uses four observable parameters, known as *Stokes parameters*, which are defined as

$$\begin{aligned} S_0 &= I_0, \\ S_1 &= I_{\mathrm{H}} - I_{\mathrm{V}}, \\ S_2 &= I_{+45} - I_{-45}, \\ S_3 &= I_{\mathrm{RCP}} - I_{\mathrm{LCP}}. \end{aligned} \tag{6.1}$$

In Eq. (6.1), I_0 represents the intensity of the given light beam, and I_{H}, I_{V}, I_{+45}, I_{-45}, I_{RCP}, and I_{LCP} represent the transmitted intensities when the given beam is passed through a linear horizontal polarizer (LHP), a linear vertical polarizer (LVP), a linear +45-deg polarizer, a linear −45-deg polarizer, a right-circular polarizer, and left-circular polarizer, respectively. Physically, S_1 gives an idea of whether the given SOP is closer to linear horizontally polarized light ($S_1 > 0$) or to linear vertically polarized light ($S_1 < 0$) or to neither ($S_1 = 0$); S_2 gives an idea of whether the given SOP is closer to the linear +45 deg ($S_2 > 0$), to linear −45 deg ($S_2 < 0$), or to neither ($S_2 = 0$); similarly, S_3 gives an idea of whether the given SOP is closer to right-circular polarized light ($S_3 > 0$), to left-circular polarized light ($S_3 < 0$), or to neither ($S_3 = 0$). Using the fact that

$$I_{\mathrm{H}} + I_{\mathrm{V}} \;=\; I_{-45} + I_{+45} \;=\; I_{\mathrm{RCP}} + I_{\mathrm{LCP}} \;=\; I_0\,, \tag{6.2}$$

Eq. (6.1) can be recast as

$$\begin{aligned} S_0 &= I_0, \\ S_1 &= 2I_{\mathrm{H}} - I_0, \\ S_2 &= 2I_{+45} - I_0, \\ S_3 &= 2I_{\mathrm{RCP}} - I_0. \end{aligned} \tag{6.3}$$

There are two other ways to define the Stokes parameters, as discussed in the following two sections.

6.2.1 In terms of amplitudes and phases of *x* and *y* components

Let the *x* and *y* linear components of the given light wave be given by

$$\begin{aligned} E_x(t) &= a_x \cos(\omega t - kz + \delta_x), \\ E_y(t) &= a_y \cos\left(\omega t - kz + \delta_y\right). \end{aligned} \tag{6.4}$$

Then, the intensity of the given wave will be given by

$$I_0 = \varepsilon_0 c\left[\left\langle {E_x}^2(t)\right\rangle + \left\langle {E_y}^2(t)\right\rangle\right] = \frac{\varepsilon_0 c}{2}\left[\left\langle {a_x}^2\right\rangle + \left\langle {a_y}^2\right\rangle\right]. \tag{6.5}$$

The intensity passed through an LHP is obviously given by

$$I_{\mathrm{H}} = \varepsilon_0 c\left[\left\langle {E_x}^2(t)\right\rangle\right] = \frac{\varepsilon_0 c}{2}\left\langle {a_x}^2\right\rangle. \tag{6.6}$$

In order to calculate the intensity passed through a linear +45-deg polarizer (L+45-deg P), we first calculate the output field, which is given by

$$E_{+45} = \frac{1}{\sqrt{2}} a_x \cos(\omega t - kz + \delta_x) + \frac{1}{\sqrt{2}} a_y \cos(\omega t - kz + \delta_y). \tag{6.7}$$

The corresponding output intensity is now given by

$$I_{+45\,\mathrm{deg}} = \varepsilon_0 c\left[\left\langle {E_{+45}}^2(t)\right\rangle\right] = \frac{\varepsilon_0 c}{2}\frac{1}{2}\left[\left\langle {a_x}^2\right\rangle + \left\langle {a_y}^2\right\rangle + \left\langle 2a_x a_y \cos\delta\right\rangle\right], \tag{6.8}$$

where $\delta = \delta_y - \delta_x$. Similarly, it can be shown that the intensity passed through a right-circular polarizer is given by

$$I_{\mathrm{RCP}} = \frac{\varepsilon_0 c}{2}\frac{1}{2}\left[\left\langle a_x{}^2\right\rangle + \left\langle a_y{}^2\right\rangle + \left\langle 2a_x a_y \sin\delta\right\rangle\right]. \tag{6.9}$$

(As discussed in Section 3.5.2, I_{RCP} can be obtained by passing the given beam through a combination of QWP with its fast axis along the *x* axis, followed by a linear analyzer with its pass axis along 45 deg with the *x* axis.)

Using Eqs. (6.5), (6.6), (6.8), and (6.9) in Eq. (6.3) after dropping the common factor ($\varepsilon_0 c/2$), one obtains the following alternative definition of the Stokes parameters:

$$\begin{aligned} S_0 &= \left\langle a_x{}^2\right\rangle + \left\langle a_y{}^2\right\rangle, \\ S_1 &= \left\langle a_x{}^2\right\rangle - \left\langle a_y{}^2\right\rangle, \\ S_2 &= \left\langle 2a_x a_y \cos\delta\right\rangle, \\ S_3 &= \left\langle 2a_x a_y \sin\delta\right\rangle. \end{aligned} \tag{6.10}$$

6.2.2 In terms of complex amplitudes of *x* and *y* components

As discussed in Chapter 5, one can also express the *x* and *y* linear components in terms of complex notations as

$$\begin{aligned} E_x(t) &= a_x e^{i(\omega t - kz + \delta_x)} = A_x e^{i(\omega t - kz)} \\ E_y(t) &= a_y e^{i(\omega t - kz + \delta_y)} = A_y e^{i(\omega t - kz)}. \end{aligned} \tag{6.11}$$

Using Eqs. (6.10) and (6.11), it is easy to show that the Stokes parameters can be expressed in terms of A_x and A_y as

$$\begin{aligned} S_0 &= \left\langle E_x E_x^*\right\rangle + \left\langle E_y E_y^*\right\rangle = \left\langle \left|A_x\right|^2\right\rangle + \left\langle \left|A_y\right|^2\right\rangle, \\ S_1 &= \left\langle E_x E_x^*\right\rangle - \left\langle E_y E_y^*\right\rangle = \left\langle \left|A_x\right|^2\right\rangle - \left\langle \left|A_y\right|^2\right\rangle, \\ S_2 &= \left\langle E_x E_y^* + E_y E_x^*\right\rangle = \left\langle 2\,\mathrm{Re}\left(A_x A_y^*\right)\right\rangle, \\ S_3 &= \left\langle i\left(E_x E_y^* - E_y E_x^*\right)\right\rangle = \left\langle 2\,\mathrm{Im}\left(A_x^* A_y\right)\right\rangle. \end{aligned} \tag{6.12}$$

In Eq. (6.12), $A_x = a_x e^{i\delta_x}$ and $A_y = a_y e^{i\delta_y}$ represent the complex amplitudes of the *x* and *y* components of a given beam.

6.3 Stokes Vectors

The Stokes parameters are generally arranged vertically in the form of a column vector or horizontally inside curly brackets in order to save space, as shown here:

$$\mathbf{S} = \begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} \quad \text{or} \quad \{S_0, S_1, S_2, S_3\}, \tag{6.13}$$

which is known as the *Stokes vector* of a given light beam. Using Eqs. (6.3), (6.10), or (6.12), the Stokes parameters of any light beam can be obtained. Generally, the Stokes vector is written in the normalized form, where all four parameters are divided by the first parameter S_0. As mentioned earlier, Stokes parameters can be used to describe completely polarized, unpolarized, as well as partially polarized light. In the following sections, we discuss some important properties of the Stokes vectors for differently polarized light beams.

6.3.1 Completely polarized light

For a completely polarized beam, a_x, a_y, and δ appearing in Eq. (6.10) are independent of time, and hence,

$$\begin{aligned} S_0 &= a_x{}^2 + a_y{}^2, \\ S_1 &= a_x{}^2 - a_y{}^2, \\ S_2 &= 2a_x a_y \cos\delta, \\ S_3 &= 2a_x a_y \sin\delta. \end{aligned} \tag{6.14}$$

In the remainder of this section, we use Eq. (6.14) to obtain the Stokes parameters for some standard states of polarization:

(i) *Linearly polarized*: For a linearly polarized beam with amplitude a and orientation θ (with the x axis), $a_x = a\cos\theta$, $a_y = a\sin\theta$, and $\delta = 0$; therefore,

$$\begin{aligned} S_0 &= a_x{}^2 + a_y{}^2 = a^2, \\ S_1 &= a_x{}^2 - a_y{}^2 = a^2\cos 2\theta, \\ S_2 &= 2a_x a_y \cos\delta = a^2 \sin 2\theta, \\ S_3 &= 2a_x a_y \sin\delta = 0. \end{aligned} \tag{6.15}$$

Thus, the Stokes vector of a linearly polarized wave oriented at θ with the x axis is given by $\{S_0, S_0\cos 2\theta, S_0 \sin 2\theta, 0\}$.

(ii) *Right-circularly polarized*: In this case,

$$a_x = a_y = a, \text{ and } \delta = \pi/2 .$$

Hence,

$$S_0 = 2a^2, \quad S_1 = S_2 = 0, \text{ and } S_3 = 2a^2 . \tag{6.16}$$

(iii) *Left-circularly polarized*: In this case,

$$a_x = a_y = a \text{ and } \delta = -\pi/2 .$$

Hence,

$$S_0 = 2a^2, \; S_1 = S_2 = 0, \text{and } S_3 = -2a^2 . \tag{6.17}$$

In view of this, the normalized Stokes vectors of the LHP ($\theta = 0$), LVP ($\theta = \pi/2$), L+45-deg P ($\theta = \pi/4$), L–45-deg P ($\theta = -\pi/4$), RCP, and LCP polarization states are given, respectively, as

LHP	LVP	L+45-deg P	L–45-deg P	RCP	LCP
$\begin{bmatrix}1\\1\\0\\0\end{bmatrix}$	$\begin{bmatrix}1\\-1\\0\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\\1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\\-1\\0\end{bmatrix}$	$\begin{bmatrix}1\\0\\0\\1\end{bmatrix}$	$\begin{bmatrix}1\\0\\0\\-1\end{bmatrix}$
↔	↕	⤢	⤡	↻	↻

(6.18)

(iv) *Elliptically polarized:* We assume that Eq. (6.4) gives the *x* and *y* components of a general elliptically polarized beam. It was shown in Section 3.3 that the orientation θ, ellipticity *b* / *a*, and sense of rotation of the polarization ellipse are given by

$$\tan 2\theta = \frac{2a_x a_y \cos\delta}{(a_x^2 - a_y^2)} \tag{6.19}$$

and

$$\sin 2\chi = \frac{2a_x a_y \sin\delta}{\left(a_x^2 + a_y^2\right)}, \tag{6.20}$$

where we have used

$$\tan\chi = \mp\frac{b}{a}. \tag{6.21}$$

In Eq. (6.21) negative and positive signs correspond to the left and right sense of rotation of the electric field vector. Further, semi-major axis *a* and semi-minor axis *b* of the polarization ellipse are given by [see Eqs. (3.38) and (3.39)]

$$a^2 = a_x^2\cos^2\theta + a_y^2\sin^2\theta + a_x a_y \sin 2\theta\cos\delta, \tag{6.22}$$

and

$$b^2 = a_x^2\sin^2\theta + a_y^2\cos^2\theta - a_x a_y \sin 2\theta\cos\delta. \tag{6.23}$$

Subtracting and adding the preceding equations, one obtains

$$a^2 - b^2 = (a_x^2 - a_y^2)\cos 2\theta + 2a_x a_y \cos\delta \sin 2\theta, \tag{6.24}$$

and

$$a^2 + b^2 = a_x^2 + a_y^2, \tag{6.25}$$

giving

$$\frac{a^2 - b^2}{a^2 + b^2} = \frac{a_x^2 - a_y^2}{a_x^2 + a_y^2}\cos 2\theta + \frac{2a_x a_y \cos\delta}{a_x^2 + a_y^2}\sin 2\theta. \tag{6.26}$$

Using Eq. (6.21), the preceding equation can be recast as

$$\cos 2\chi = \frac{a_x^2 - a_y^2}{a_x^2 + a_y^2}\cos 2\theta + \frac{2a_x a_y \cos\delta}{a_x^2 + a_y^2}\sin 2\theta. \tag{6.27}$$

Using Eq. (6.14), Eqs. (6.19), (6.20), and (6.27) can be written as

$$\tan 2\theta = \frac{S_2}{S_1}, \tag{6.28}$$

$$\sin 2\chi = \frac{S_3}{S_0}, \tag{6.29}$$

and

$$\begin{aligned}\cos 2\chi &= \frac{S_1}{S_0}\cos 2\theta + \frac{S_2}{S_0}\sin 2\theta \\ &= \frac{S_1}{S_0}\left[\cos 2\theta + \frac{S_2}{S_1}\sin 2\theta\right] = \frac{S_1}{S_0}\frac{1}{\cos 2\theta}.\end{aligned} \tag{6.30}$$

Using these three equations, one can easily show that the Stokes parameters of a completely polarized beam are given as

$$\begin{aligned}S_1 &= S_0 \cos 2\chi \cos 2\theta, \\ S_2 &= S_0 \cos 2\chi \sin 2\theta, \\ S_3 &= S_0 \sin 2\chi.\end{aligned} \tag{6.31}$$

The preceding relations show that for completely polarized light, the four Stokes parameters are not independent and are related to each other through the following relation:

$$S_0^{\,2} = S_1^{\,2} + S_2^{\,2} + S_3^{\,2}. \tag{6.32}$$

Thus, the normalized Stokes vector of a completely polarized beam is

$$\mathbf{S} = \begin{bmatrix} 1 \\ \cos 2\chi \cos 2\theta \\ \cos 2\chi \sin 2\theta \\ \sin 2\chi \end{bmatrix}. \tag{6.33}$$

Conversely, if the Stokes parameters of a given polarized wave are known, the orientation θ and the ellipticity angle χ of its polarization ellipse can be obtained using the following relations:

$$2\theta = \tan^{-1}\left(\frac{S_2}{S_1}\right), \quad 0 \le \theta \le \pi,$$

and

$$2\chi = \sin^{-1}\left(\frac{S_3}{S_0}\right), \quad -\frac{\pi}{4} \le \chi \le \frac{\pi}{4}. \tag{6.34}$$

6.3.2 Unpolarized light

In the case of unpolarized light, $a_x = a_y = a$, and δ is a randomly varying parameter with time. As a result, $\langle \cos\delta \rangle = \langle \sin\delta \rangle = 0$. This means that

$$S_0 = a_x^2 + a_y^2 = 2a^2 \quad \text{and} \quad S_1 = S_2 = S_3 = 0. \tag{6.35}$$

Thus, the corresponding normalized Stokes vector is given by

$$\mathbf{S} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{6.36}$$

It may be noted that in this case, $S_0^{\,2} \neq S_1^{\,2} + S_2^{\,2} + S_3^{\,2}$. Instead, $S_0^{\,2} > S_1^{\,2} + S_2^{\,2} + S_3^{\,2}$, which is also the case for partially polarized light, as discussed in the next section.

6.3.3 Partially polarized light

Partially polarized light is a mixture of polarized and unpolarized light, and its Stokes vector can be expressed as a sum of its polarized and unpolarized components as

$$\mathbf{S}_{\text{par}} = \mathbf{S}_{\text{pol}} + \mathbf{S}_{\text{unp}}. \tag{6.37}$$

In matrix form, Eq. (6.37) can be expressed as

$$\begin{bmatrix} S_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} = \begin{bmatrix} pS_0 \\ S_1 \\ S_2 \\ S_3 \end{bmatrix} + \begin{bmatrix} (1-p)S_0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{6.38}$$

where *p*, known as the degree of polarization (DOP) of the given partially polarized light, represents the fraction of the light beam intensity that is completely polarized. Thus, S_0, pS_0, and $(1-p)S_0$, represent the intensities of the given beam (partially polarized), its polarized component, and its unpolarized component, respectively. The first Stokes vector on the right-hand side in Eq. (6.38) represents a completely polarized beam:

$$(pS_0)^2 = S_1^{\,2} + S_2^{\,2} + S_3^{\,2}, \tag{6.39}$$

i.e.,

$$p = \frac{\sqrt{S_1^{\,2} + S_2^{\,2} + S_3^{\,2}}}{S_0} = \frac{I_{\text{pol}}}{I_{\text{tot}}}\ , \tag{6.40}$$

where I_{pol} corresponds to the intensity of the polarized component, and I_{tot} is the total intensity of the given partially polarized beam. It may be noted that $p = 0$ represents the unpolarized light, $p = 1$ represents the polarized light, and $0 < p < 1$ represents the partially polarized light.

Let θ and χ represent the orientation and the ellipticity angle of the polarized component. Since pS_0 represents its intensity, its first Stokes parameter will be pS_0, and the other three Stokes parameters will be given by

$$\begin{aligned} S_1 &= pS_0 \cos 2\chi \cos 2\theta, \\ S_2 &= pS_0 \cos 2\chi \sin 2\theta, \\ S_3 &= pS_0 \sin 2\chi. \end{aligned} \tag{6.41}$$

If the Stokes vector of the partially polarized beam is in the normalized form, then $S_0 = 1$, and $p = \sqrt{S_1^{\,2} + S_2^{\,2} + S_3^{\,2}}$. In this case, the Stokes vector of the polarized component will be

$$\begin{aligned} S_1 &= p \cos 2\chi \cos 2\theta, \\ S_2 &= p \cos 2\chi \sin 2\theta, \\ S_3 &= p \sin 2\chi. \end{aligned} \tag{6.42}$$

Equation (6.42) suggests that in the Stokes subspace (S_1, S_2, S_3), the SOP of any light wave can be described by a point on a sphere of radius p with Cartesian coordinates as (S_1, S_2, S_3) and polar coordinates as (p, 2θ, $\pi/2 - 2\chi$) as shown in Fig. 6.1. This is an important property of the Stokes parameters that is used quite frequently to represent the Stokes vector as a 3D vector in Stokes subspace (S_1, S_2, S_3). In the case of completely polarized light, the previously mentioned sphere is nothing but the Poincaré sphere, which is discussed in the next chapter.

Example 6.1 Consider partially polarized light whose Stokes vector is given by $\{5,\ 4,\ 0,\ 0\}$. (i) Obtain its DOP and (ii) write the given Stokes vector as a sum of its polarized and unpolarized components.

Solution: The DOP is given by [Eq. (6.40)]:

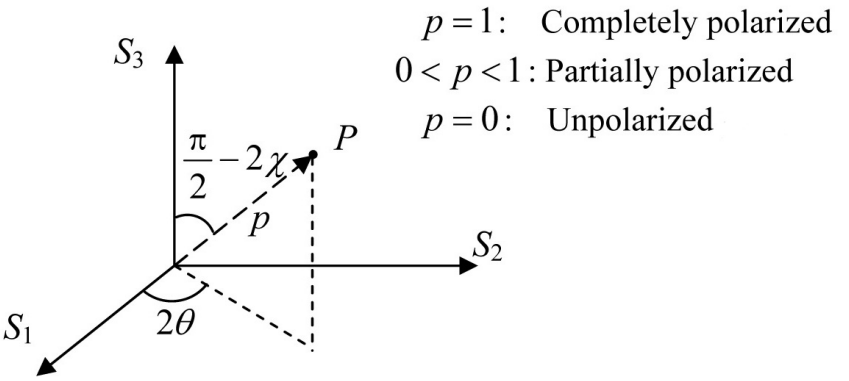


Figure 6.1 Schematic representation of a general SOP in Stokes subspace.

$$p = \frac{\sqrt{S_1^{\,2} + S_2^{\,2} + S_3^{\,2}}}{S_0} = \frac{4}{5} = 0.8\,.$$

Thus, the given Stokes vector can be written as [Eq. (6.38)]

$$\begin{bmatrix} 5 \\ 4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.8\times 5 \\ 4 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0.2\times 5 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 4\begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

which means that the intensity of the given beam is 5 units, out of which 80% is linear horizontally polarized and 20% is unpolarized.

6.4 Determination of Stokes Vectors

As suggested by Eq. (6.3), the Stokes vector of a given light beam can be determined first by measuring its intensity directly and then measuring its intensity after passing it through three different polarizers—namely, a linear horizontal, a linear +45-deg, and a right-circular polarizer. The measured intensities will be $I_0, I_{\mathrm{H}}, I_{+45}$, and I_{RCP} respectively, from which the Stokes parameters S_0, S_1, S_2 and S_3 can easily be calculated using Eq. (6.3). We demonstrate this using the following simple example.

Example 6.2 Determine the Stokes parameters of a linearly polarized light oriented at an angle θ with the *x* axis.

Solution: Let the intensity of the given beam be I_0. It can easily be shown that if this beam is passed separately through a linear horizontal polarizer, a linear +45-

deg polarizer, and a right-circular polarizer, the intensities I_{H}, I_{+45}, and I_{RCP} of the output beam obtained in the three cases will be given by

$$\begin{aligned} I_{\mathrm{H}} &= I_0 \cos^2\theta, \\ I_{+45} &= I_0 \cos^2(\pi/4-\theta) = \frac{I_0}{2}(1+\sin 2\theta), \\ I_{\mathrm{RCP}} &= \frac{I_0}{2}. \end{aligned} \tag{6.43}$$

Using Eq. (6.3), the Stokes parameters will be given as

$$\begin{aligned} S_0 &= I_0, \\ S_1 &= 2I_{\mathrm{H}} - I_0 = 2I_0\cos^2\theta - I_0 = I_0\cos 2\theta, \\ S_2 &= 2I_{+45} - I_0 = I_0(1+\sin 2\theta) - I_0 = I_0 \sin 2\theta, \\ S_3 &= 2I_{\mathrm{RCP}} - I_0 = I_0 - I_0 = 0. \end{aligned} \tag{6.44}$$

Thus, the normalized Stokes vector of a linearly polarized light oriented at an angle θ with the *x* axis is given by

$$S = \begin{bmatrix} 1 \\ \cos 2\theta \\ \sin 2\theta \\ 0 \end{bmatrix}, \tag{6.45}$$

which is the normalized form of Eq. (6.15). Similarly, Stokes parameters of other polarization states can also be determined.

6.5 Mueller Matrices

Like Jones vectors, Stokes vectors of the input and output states of polarization corresponding to a polarizing device can also be related by a square matrix (4 × 4 in this case), which is known as a *Mueller matrix*, after Hans Mueller, who devised this method in 1943. According to this method, if $\mathbf{S}$ and $\mathbf{S}'$ are the Stokes vectors of the input and output states of polarization of a polarizing device, they are related as

$$\mathbf{S}' = M\,\mathbf{S}, \tag{6.46}$$

where *M*, known as the Mueller matrix, is a 4 × 4 matrix of the form

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix}. \tag{6.47}$$

6.6 Determination of Mueller Matrices

The Mueller matrix of any optical device can be obtained by passing four light beams of different polarization states—namely, an unpolarized, a linear horizontal, a linear +45-deg, and a right-circularly polarized—through the given device and then determining the Stokes parameters of the output beam in each case.

Let the Mueller matrix of the given device be denoted by *M* and given by Eq. (6.47). We first pass an unpolarized beam {1, 0, 0, 0} through the given device. The Stokes vector of the output beam will then be given by

$$\mathbf{S}_0' = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} m_{11} \\ m_{21} \\ m_{31} \\ m_{41} \end{bmatrix}. \tag{6.48}$$

The matrix elements $m_{11}, m_{21}, m_{31}, \text{and } m_{41}$ are equal to the Stokes parameters of the output beam and are thus determined. We now pass a linear horizontally polarized {1, 1, 0, 0} light beam through the device. The Stokes vector of the output beam will then be given by

$$\mathbf{S}_1' = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} m_{11} + m_{12} \\ m_{21} + m_{22} \\ m_{31} + m_{32} \\ m_{41} + m_{42} \end{bmatrix}. \tag{6.49}$$

In this case, the Stokes parameters of the output beam are equal to the sum of the matrix elements in the first two columns of the Mueller matrix. Having determined the elements of the first column in the previous case, one can easily obtain $m_{12}, m_{22}, m_{32}, \text{ and } m_{42}$ by measuring the Stokes vector of the output SOP.

Similarly, if we pass a linear +45-deg polarized {1, 0, 1, 0} beam and a right-circularly polarized {1, 0, 0, 1} beam through the device, the Stokes vectors of the corresponding output beams will be given by

$$\mathbf{S}_2' = \begin{bmatrix} m_{11}+m_{13} \\ m_{21}+m_{23} \\ m_{31}+m_{33} \\ m_{41}+m_{43} \end{bmatrix} \quad \text{and} \quad \mathbf{S}_3' = \begin{bmatrix} m_{11}+m_{14} \\ m_{21}+m_{24} \\ m_{31}+m_{34} \\ m_{41}+m_{44} \end{bmatrix}. \tag{6.50}$$

By determining the preceding Stokes parameters and using the elements $m_{11}, m_{21}, m_{31}, \text{and}\, m_{41}$ already determined, one can determine the remainder of the elements. Using the preceding method, in the following section, we determine the Mueller matrices of some basic polarization components.

6.6.1 Mueller matrix of a linear polarizer

We consider a linear polarizer whose pass axis makes an angle θ with the x axis. As discussed earlier, we will pass an unpolarized, a linear horizontal, a linear +45-deg, and a right-circularly polarized light beam individually through the given device and then determine the Stokes parameters of the output beam in each case. Let the intensity of the input beam in each case be unity. It should be noted that in each case, the SOP of the output beam is linear, making an angle θ with the x axis whose normalized Stokes vector is given by Eq. (6.45) as

$$\mathbf{S} = \begin{bmatrix} 1 \\ \cos 2\theta \\ \sin 2\theta \\ 0 \end{bmatrix}. \tag{6.51}$$

The intensities of the output beams, however, will be different in the different cases discussed earlier. In the first case (unpolarized beam), the output intensity will be 1/2, and hence the Stokes vector of the output beam will be

$$\mathbf{S}_0' = \frac{1}{2}\begin{bmatrix} 1 \\ \cos 2\theta \\ \sin 2\theta \\ 0 \end{bmatrix} = \begin{bmatrix} m_{11} \\ m_{21} \\ m_{31} \\ m_{41} \end{bmatrix}. \tag{6.52}$$

This means that

$$m_{11} = \frac{1}{2},\ m_{21} = \frac{1}{2}\cos 2\theta,\ m_{31} = \frac{1}{2}\sin 2\theta,\ \text{and}\ m_{41} = 0. \tag{6.53}$$

In the second case, using a linear horizontal input beam, the output intensity will be $\cos^2\theta$ (using Malus' law), and hence the corresponding Stokes vector will be

$$\mathbf{S}_1' = \cos^2\theta \begin{bmatrix} 1 \\ \cos 2\theta \\ \sin 2\theta \\ 0 \end{bmatrix} = \begin{bmatrix} m_{11} + m_{12} \\ m_{21} + m_{22} \\ m_{31} + m_{32} \\ m_{41} + m_{42} \end{bmatrix}. \tag{6.54}$$

Subtracting Eq. (6.52) from Eq. (6.54), one obtains the other four matrix elements as

$$\begin{aligned} m_{12} &= \cos^2\theta - \frac{1}{2} = \frac{1}{2}\cos 2\theta, \\ m_{22} &= \cos^2\theta\cos 2\theta - \frac{1}{2}\cos 2\theta = \frac{1}{2}\cos^2 2\theta, \\ m_{32} &= \cos^2\theta\sin 2\theta - \frac{1}{2}\sin 2\theta = \frac{1}{2}\sin 2\theta\cos 2\theta, \\ m_{42} &= 0. \end{aligned} \tag{6.55}$$

In the third case of a linear +45-deg input beam, the output intensity will be $\cos^2(\pi/4 - \theta)$, i.e., $1/2\,(1 + \sin 2\theta)$, and hence the corresponding normalized Stokes vector will be

$$\mathbf{S}_2' = \frac{1}{2}(1+\sin 2\theta) \begin{bmatrix} 1 \\ \cos 2\theta \\ \sin 2\theta \\ 0 \end{bmatrix} = \begin{bmatrix} m_{11} + m_{13} \\ m_{21} + m_{23} \\ m_{31} + m_{33} \\ m_{41} + m_{43} \end{bmatrix}. \tag{6.56}$$

Subtracting Eq. (6.52) from Eqs. (6.56), one obtains

$$\begin{aligned} m_{13} &= \frac{1}{2}\sin 2\theta, \qquad & m_{23} &= \frac{1}{2}\sin 2\theta\cos 2\theta, \\ m_{33} &= \frac{1}{2}\sin^2 2\theta, \qquad & m_{43} &= 0. \end{aligned} \tag{6.57}$$

In the last case of the right-circularly polarized input beam, the output intensity will be 1/2, and hence the corresponding normalized Stokes vector will be

$$\mathbf{S}_3' = \frac{1}{2}\begin{bmatrix} 1 \\ \cos 2\theta \\ \sin 2\theta \\ 0 \end{bmatrix} = \begin{bmatrix} m_{11} + m_{14} \\ m_{21} + m_{24} \\ m_{31} + m_{34} \\ m_{41} + m_{44} \end{bmatrix}. \tag{6.58}$$

Subtracting Eq. (6.52) from Eq (6.58), one obtains the remainder of the matrix elements as

$$m_{14} = \; m_{24} = \; m_{34} = \; m_{44} = 0\,. \tag{6.59}$$

Thus, the Mueller matrix of a linear polarizer with its pass axis at an angle θ with the *x* axis is given as

$$M_{\mathrm{LP}} = \frac{1}{2}\begin{bmatrix} 1 & \cos 2\theta & \sin 2\theta & 0 \\ \cos 2\theta & \cos^2 2\theta & \sin 2\theta \cos 2\theta & 0 \\ \sin 2\theta & \sin 2\theta \cos 2\theta & \sin^2 2\theta & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.60}$$

Substituting θ = 0, 90 deg, +45 deg, and –45 deg in the preceding equation, one obtains Mueller matrices for linear horizontal, linear vertical, +45-deg, and −45-deg polarizers, which are listed here:

LHP: $$\frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.61}$$

LVP: $$\frac{1}{2}\begin{bmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.62}$$

Linear +45-deg polarizer: $$\frac{1}{2}\begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.63}$$

Linear −45-deg polarizer: $$\frac{1}{2}\begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.64}$$

6.6.2 Mueller matrix of a circular polarizer

Once again, we pass an unpolarized, a linear horizontal, a linear +45-deg, and a right-circularly polarized beam, each of unit intensity, from the given device. Obviously, in all of the cases, the output polarization state will be right circular; however, in the first three cases, the output intensity will be 1/2, while in the last case, it will be 1. Using this information and following the procedure discussed earlier, one can easily obtain the Mueller matrix of an ideal right-circular polarizer, which is given as

Ideal right-circular polarizer: $$\frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}. \tag{6.65}$$

Similarly, one can obtain the Mueller matrix of an ideal left-circular polarizer, which is given as

Ideal left-circular polarizer: $$\frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}. \tag{6.66}$$

6.6.3 Mueller matrix of a linear retarder

Once again, we pass an unpolarized, linear horizontal, linear + 45-deg, and right-circularly polarized beam, each of unit intensity, from the given device. It should be noted that the output intensity in each case will be unity, as an ideal retarder does not absorb any energy. The Stokes vector of the output SOP can be obtained in each case as follows:

Case I: If the incident beam is an unpolarized beam, obviously the output will also be unpolarized. Thus, the Stokes vectors of the input as well as of the output SOP will be

$$\mathbf{S_{out}} = \mathbf{S_{in}} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \Rightarrow \begin{matrix} m_{11} = 1 \\ m_{21} = 0 \\ m_{31} = 0 \\ m_{41} = 0. \end{matrix} \tag{6.67}$$

In order to obtain the Stokes vector in the remaining three cases, we assume that the fast axis of the retarder is at an angle θ with the x axis, as shown in Fig. 6.2. In each case, the incident SOP will excite fast and slow components, for instance, of amplitudes a_{f} and a_{s}, which will propagate with different phase velocities.

Let δ be the phase difference introduced between the fast and slow components. The complex amplitudes of the x and y components of the output SOP will then be given by

$$A_x = a_{\mathrm{f}} e^{i\delta} \cos\theta - a_{\mathrm{s}} \sin\theta, \tag{6.68a}$$

$$A_y = a_{\mathrm{f}} e^{i\delta} \sin\theta + a_{\mathrm{s}} \cos\theta. \tag{6.68b}$$

For the remaining three cases, we give the values of a_{f} and a_{s} and obtain the Stokes vector of the output SOP in each case, determining the various elements of the Mueller matrix.

Case II: If the input SOP is linear horizontal, then

$$a_{\mathrm{f}} = \cos\theta \text{ and } a_{\mathrm{s}} = -\sin\theta.$$

Substituting a_{f} and a_{s} in Eqs. (6.68a) and (6.68b), we obtain

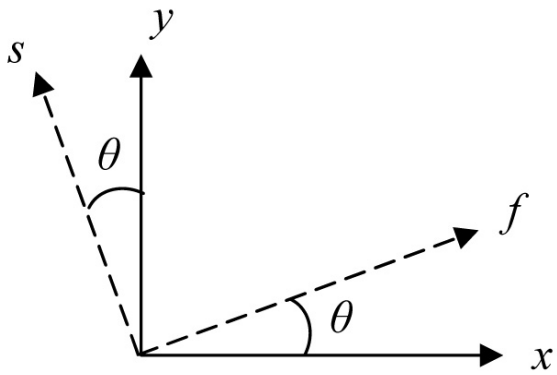


Figure 6.2 Schematic of a linear retarder with its fast axis at angle θ with the x axis.

$$A_x = e^{i\delta}\cos^2\theta + \sin^2\theta,$$
$$A_y = (e^{i\delta} - 1)\sin\theta\cos\theta. \quad (6.69)$$

Substituting the preceding values of the complex amplitudes in Eq. (6.12), the Stokes vector of the output SOP is

$$\mathbf{S}_{\text{out}} = \begin{bmatrix} 1 \\ \cos^2 2\theta + \sin^2 2\theta\cos\delta \\ (1-\cos\delta)\sin 2\theta\cos 2\theta \\ \sin 2\theta\sin\delta \end{bmatrix} \Rightarrow \begin{bmatrix} m_{12} \\ m_{22} \\ m_{32} \\ m_{42} \end{bmatrix} = \begin{bmatrix} 0 \\ \cos^2 2\theta + \sin^2 2\theta\cos\delta \\ (1-\cos\delta)\sin 2\theta\cos 2\theta \\ \sin 2\theta\sin\delta \end{bmatrix}. \quad (6.70)$$

Case III: If the input SOP is linear +45 deg, then

$$a_{\text{f}} = \cos(\pi/4 - \theta) \text{ and } a_{\text{s}} = \sin(\pi/4 - \theta).$$

Substituting the preceding values in Eqs. (6.68a) and (6.68b), one obtains

$$A_x = e^{i\delta}\cos(\pi/4-\theta)\cos\theta - \sin(\pi/4-\theta)\sin\theta,$$
$$A_y = e^{i\delta}\cos(\pi/4-\theta)\sin\theta + \sin(\pi/4-\theta)\cos\theta. \quad (6.71)$$

Substituting A_x and A_y in Eq. (6.12), the Stokes vector of the output SOP is

$$\mathbf{S}_{\text{out}} = \begin{bmatrix} 1 \\ (1-\cos\delta)\sin 2\theta\cos 2\theta \\ \sin^2 2\theta + \cos^2 2\theta\cos\delta \\ -\cos 2\theta\sin\delta \end{bmatrix} \Rightarrow \begin{bmatrix} m_{13} \\ m_{23} \\ m_{33} \\ m_{43} \end{bmatrix} = \begin{bmatrix} 0 \\ (1-\cos\delta)\sin 2\theta\cos 2\theta \\ \sin^2 2\theta + \cos^2 2\theta\cos\delta \\ -\cos 2\theta\sin\delta \end{bmatrix}. \quad (6.72)$$

Case IV: If the input SOP is right-circularly polarized, then

$$a_{\text{f}} = \frac{1}{\sqrt{2}} \quad \text{and} \quad a_{\text{s}} = \frac{1}{\sqrt{2}}e^{\frac{i\pi}{2}} = \frac{i}{\sqrt{2}}.$$

Substituting the values in Eqs. (6.68a) and (6.68b), one obtains

$$A_x = \frac{1}{\sqrt{2}}(e^{i\delta}\cos\theta - i\sin\theta),$$
$$A_y = \frac{1}{\sqrt{2}}(e^{i\delta}\sin\theta + i\cos\theta). \quad (6.73)$$

Using the values of the complex amplitudes in Eq. (6.12), the Stokes vector of the output SOP is

$$\mathbf{S}_{\mathbf{out}} = \begin{bmatrix} 1 \\ -\sin 2\theta \sin\delta \\ \cos 2\theta \sin\delta \\ \cos\delta \end{bmatrix} \Rightarrow \begin{bmatrix} m_{14} \\ m_{24} \\ m_{34} \\ m_{44} \end{bmatrix} = \begin{bmatrix} 0 \\ -\sin 2\theta \sin\delta \\ \cos 2\theta \sin\delta \\ \cos\delta \end{bmatrix}. \tag{6.74}$$

Having determined all of the elements, the Mueller matrix of a linear retarder with retardation δ and its fast axis at an angle θ with the x axis is given by

$$M_{\text{ret}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos^2 2\theta + \sin^2 2\theta \cos\delta & (1-\cos\delta)\sin 2\theta \cos 2\theta & -\sin 2\theta \sin\delta \\ 0 & (1-\cos\delta)\sin 2\theta \cos 2\theta & \sin^2 2\theta + \cos^2 2\theta \cos\delta & \cos 2\theta \sin\delta \\ 0 & \sin 2\theta \sin\delta & -\cos 2\theta \sin\delta & \cos\delta \end{bmatrix}. \tag{6.75}$$

Using Eq. (6.75), one can easily obtain the Mueller matrices for a QWP ($\delta = \pi/2$) and an HWP ($\delta = \pi$), as given here:

QWP, with its fast axis along the x axis ($\theta = 0$, $\delta = \pi/2$):

$$M_{\text{QWP}}(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}. \tag{6.76}$$

QWP, with its fast axis along the y axis ($\theta = \pi/2$, $\delta = \pi/2$):

$$M_{\text{QWP}}(\pi/2) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{6.77}$$

HWP, with its fast axis along the x axis ($\theta = 0$) or y axis ($\theta = \pi/2$):

$$M_{\mathrm{HWP}}(0)=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{bmatrix}. \tag{6.78}$$

In the following section, we provide some examples of the use and construction of the Mueller matrices.

6.6.4 Mueller matrix of a rotator

A *rotator* is a device that rotates the polarization ellipse of an incident SOP by an angle δ without affecting its ellipticity or the sense of rotation. In order to obtain the various elements of its Mueller matrix, we follow the same procedure, as discussed here:

Step 1: We pass an unpolarized beam {1, 0, 0, 0} through the rotator. Obviously, the output beam will also be unpolarized and represented by {1, 0, 0, 0}. This means that

$$\begin{bmatrix}m_{11}&m_{12}&m_{13}&m_{14}\\m_{21}&m_{22}&m_{23}&m_{24}\\m_{31}&m_{32}&m_{33}&m_{34}\\m_{41}&m_{42}&m_{43}&m_{44}\end{bmatrix}\begin{bmatrix}1\\0\\0\\0\end{bmatrix}=\begin{bmatrix}1\\0\\0\\0\end{bmatrix}\quad\Rightarrow\quad\begin{aligned}m_{11}&=1\\m_{21}&=0\\m_{31}&=0\\m_{41}&=0.\end{aligned} \tag{6.79}$$

Step 2: We now pass a linear horizontal polarization state {1, 1, 0, 0} through the rotator. In this case, the output will be rotated linearly polarized, oriented along δ, and thus will be given by $\{1, \cos 2\delta, \sin 2\delta, 0\}$. This means that

$$\begin{bmatrix}1&m_{12}&m_{13}&m_{14}\\0&m_{22}&m_{23}&m_{24}\\0&m_{32}&m_{33}&m_{34}\\0&m_{42}&m_{43}&m_{44}\end{bmatrix}\begin{bmatrix}1\\1\\0\\0\end{bmatrix}=\begin{bmatrix}1\\\cos 2\delta\\\sin 2\delta\\0\end{bmatrix}\quad\Rightarrow\quad\begin{aligned}m_{12}&=0\\m_{22}&=\cos 2\delta\\m_{32}&=\sin 2\delta\\m_{42}&=0.\end{aligned} \tag{6.80}$$

Step 3: Next, we pass a linear +45-deg polarization state {1, 0, 1, 0} through the rotator. In this case, the output will be rotated linearly polarized, oriented along $\delta + \pi/4$, and thus will be given by {1, –sin 2δ, cos 2δ, 0}. This means that

$$\begin{bmatrix} 1 & m_{12} & m_{13} & m_{14} \\ 0 & m_{22} & m_{23} & m_{24} \\ 0 & m_{32} & m_{33} & m_{34} \\ 0 & m_{42} & m_{43} & m_{44} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -\sin 2\delta \\ \cos 2\delta \\ 0 \end{bmatrix} \quad \Rightarrow \quad \begin{aligned} m_{13} &= 0 \\ m_{23} &= -\sin 2\delta \\ m_{33} &= \cos 2\delta \\ m_{43} &= 0. \end{aligned} \tag{6.81}$$

Step 4: Last, we pass a right-circular polarization state {1, 0, 0, 1} through the rotator. Obviously, the output will remain unchanged and thus will be given by {1, 0, 0, 1} again. This means that

$$\begin{bmatrix} 1 & m_{12} & m_{13} & m_{14} \\ 0 & m_{22} & m_{23} & m_{24} \\ 0 & m_{32} & m_{33} & m_{34} \\ 0 & m_{42} & m_{43} & m_{44} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad \Rightarrow \quad \begin{aligned} m_{14} &= 0 \\ m_{24} &= 0 \\ m_{34} &= 0 \\ m_{44} &= 1. \end{aligned} \tag{6.82}$$

Now, the elements of the entire matrix are known, and the desired Mueller matrix of a rotator is given by

$$M_{\text{rot}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\delta & -\sin 2\delta & 0 \\ 0 & \sin 2\delta & \cos 2\delta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{6.83}$$

It should be noted that a circularly birefringent medium also works as a rotator (see Example 5.8). Thus, the Mueller matrix in Eq. (6.83) also corresponds to a circularly birefringent medium introducing a phase difference of 2δ between the left- and right-circular components propagating through the medium, with the left-circular component as the fast component.

Example 6.3 Consider partially polarized light whose Stokes vector is given by {5, 4, 0, 0}. Obtain the output SOP and its intensity if it is passed through an LHP.

Solution: The Mueller matrix of the LHP is given by Eq. (6.61). The Stokes vector of the output SOP will be

$$\frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ 4 \\ 0 \\ 0 \end{bmatrix} = 4.5 \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Thus, the output SOP is a linear horizontally polarized wave whose intensity is 4.5 units (i.e., 90% of the input).

Example 6.4 Verify that the rotator described by Eq. (6.83) rotates only the polarization ellipse of a given SOP passing through it without affecting its ellipticity or the sense of rotation.

Solution: We consider that a general elliptically polarized wave, described by

$$S_{\text{in}} = \begin{bmatrix} 1 \\ \cos 2\chi \cos 2\theta \\ \cos 2\chi \sin 2\theta \\ \sin 2\chi \end{bmatrix},$$

is passed through a rotator described by the Mueller matrix given by Eq. (6.83). The Stokes vector of the output will be

$$S_{\text{out}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\delta & -\sin 2\delta & 0 \\ 0 & \sin 2\delta & \cos 2\delta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ \cos 2\chi \cos 2\theta \\ \cos 2\chi \sin 2\theta \\ \sin 2\chi \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ \cos 2\chi \cos\{2\theta + 2\delta\} \\ \cos 2\chi \sin\{2\theta + 2\delta\} \\ \sin 2\chi \end{bmatrix},$$

which represents an elliptically polarized beam that is the same as the input beam except that the orientation is increased by δ.

Example 6.5 Show that a combination of two HWPs whose fast axes are rotated by $\pi/4$ with respect to each other act as a rotator, rotating the polarization ellipse by $\pi/2$.

Solution: Let the fast axes of the two HWPs be oriented along $\theta = 0$ and $\theta = \pi/4$, respectively, the Mueller matrices of which can be obtained from Eq. (6.75) ($\delta = \pi$). The Mueller matrix of the combination will then be given by

$$M = M_{\mathrm{HWP}}(\pi/4)M_{\mathrm{HWP}}(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}.$$

Therefore, $$M = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

which represents a rotator with $\delta = \pi/2$ [see Eq. (6.83)].

Bibliography

Azzam, R. M. A. and N. M. Bashara, *Ellipsometry and Polarized Light*, North Holland Publishing Company, Amsterdam (1977).

Collett, E., *Field Guide to Polarization*, SPIE Press, Bellingham, WA (2005). [doi: 10.1117/3.626141].

Goldstein, D., *Polarized Light*, 2nd ed., Marcel Dekker, New York (2003).

Shurcliff, W. A., *Polarized Light: Production and Use*, Harvard University Press, Cambridge, MA (1962).

Stokes, G. G., "On the composition and resolution of streams of polarized light from different sources," *Trans. Cambridge Phil. Soc.* **9**, 399–416 (1852).

Theocaris, P. S. and E. E. Gdoutos, *Matrix Theory of Photoelasticity*, Springer Verlag, Berlin–Heidelberg (1979).

Chapter 7
Poincaré Sphere Representation of Polarized Light

7.1 Introduction

The Poincaré sphere representation was conceived by the French physicist Henri Poincaré in 1892. It is a simple and extremely useful geometrical representation of various polarization states and their evolution through a birefringent medium. We will see that this representation makes the formulation as well as the explanation of a difficult problem very easy due to the graphical depiction of the actions of the polarizers and birefringent media involved. According to this representation, one represents various polarization states, polarizers, and birefringent media by specific points on the surface of a sphere (unit radius) in terms of the longitude and latitude, as discussed in the following section.

7.2 Various Polarization States

The most general polarization state is an elliptically polarized state, which is described by the following three parameters:

(i) The orientation θ of the major axis of the polarization ellipse, which is measured with respect to some fixed direction. Let us select this direction as the horizontal direction (x axis) in the (x-y) plane, transverse to the direction of propagation, taken as the z axis (see Fig. 7.1).

(ii) The ellipticity of the ellipse representing the SOP, which is measured in terms of an angle $\varepsilon = \tan^{-1}(b/a)$, where b and a represent the semi-minor and semi-major axess, respectively. Obviously, a and b are both positive quantities, and $a \geq b$; thus $0 \leq \varepsilon \leq \pi/4$.

(iii) The sense of rotation of the electric field with time is specified by assigning a positive or negative sign to ε; the positive sign is for right rotations, and the negative sign for left rotations. Here, the right/left rotations are the clockwise/counterclockwise rotations, as seen by an observer looking into the source of the light.

Thus, a general SOP can be uniquely defined in terms of two angles—namely, θ and $\chi = \pm\varepsilon$, called the azimuth and ellipticity angle, respectively. All possible polarization states are covered by the following range of θ and χ:

$$0 \le \theta \le \pi \text{ and } -\frac{\pi}{4} \le \chi \le +\frac{\pi}{4}. \tag{7.1}$$

Figure 7.1 shows the polarization ellipse for a right- and left-elliptic polarization state with orientation $\theta = \pi/4$ and ellipticity $b/a = 1/\sqrt{3}$, i.e., $\varepsilon = \pi/6$.

7.3 Poincaré Sphere Representation

A given SOP can be uniquely represented by a point on the surface of a sphere of unit radius, whose longitude and latitude have the values 2θ and 2χ, respectively (see Fig. 7.2). This representation is called the *Poincaré sphere representation*, and the sphere is called the *Poincaré sphere* (PS).

Figure 7.3 enlists some of the standard polarization states and their representation on the PS. Some general rules of representing various polarization states on the PS are given here:

(i) For a linearly polarized state, $b/a = 0$, and hence, $\chi = 0$. Thus, various linearly polarized states are represented by various points on the equator ($\chi = 0$). Let point H represent the linear horizontal polarization whose parameters are (0, 0), and then the diametrically opposite point V (π, 0) represents the linear vertical SOP.

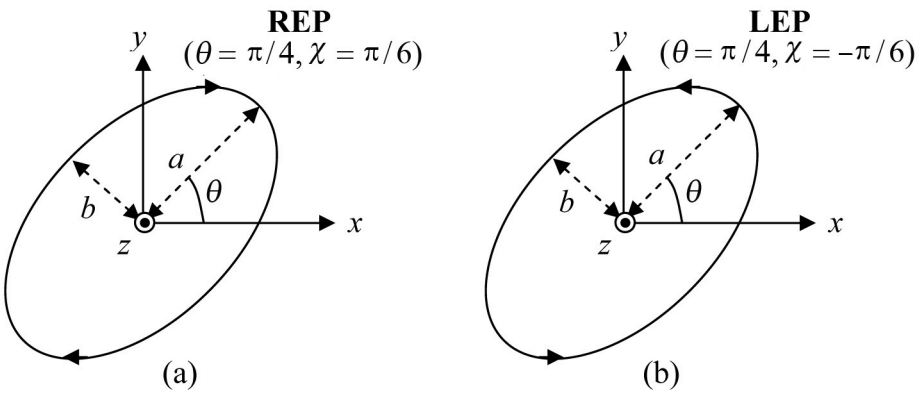


Figure 7.1 Polarization ellipse of (a) a right- and (b) a left-elliptic SOP with $\theta = \pi/4$. The propagation is along the *z* direction, out of the page. We have assumed that $a = \sqrt{3}b$.

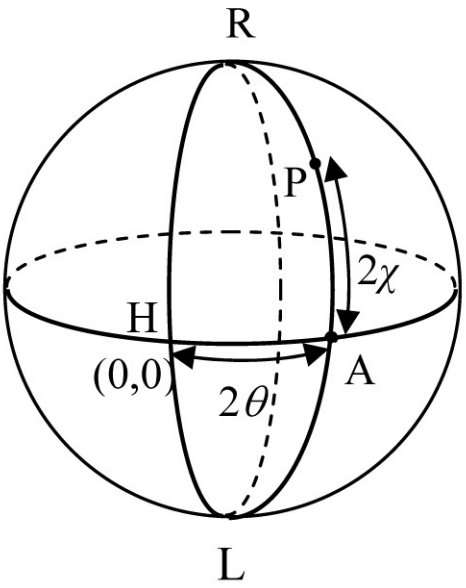


Figure 7.2 Poincaré sphere representation of a general SOP.

(ii) All right-rotating polarization states have positive values of χ and hence lie on the upper hemisphere; all left-rotating SOPs lie on the lower hemisphere.

(iii) For circularly polarized light, $b/a = 1$, and hence, $\varepsilon = \pi/4$. Thus, the right- and left-circular polarization states are characterized by

RCP: $2\chi = +\pi/2$,

LCP: $= -\pi/2$.

On the Poincaré sphere, north and south poles correspond to $2\chi = +\pi/2$ and $-\pi/2$, respectively, and represent the right- (R) and left- (L) circular SOP. Orientation θ has no meaning for circularly polarized light.

(iv) Points on a circle parallel to the equator, i.e., a latitude circle [see Fig. 7.3(a)], have the same values of χ and hence represent various elliptic SOPs with the same ellipticity but different orientations.

(v) Points on the meridian RAL [see Fig. 7.3(a)] correspond to a constant value of θ and hence represent different elliptic polarization states with the same orientation θ but different ellipticities.

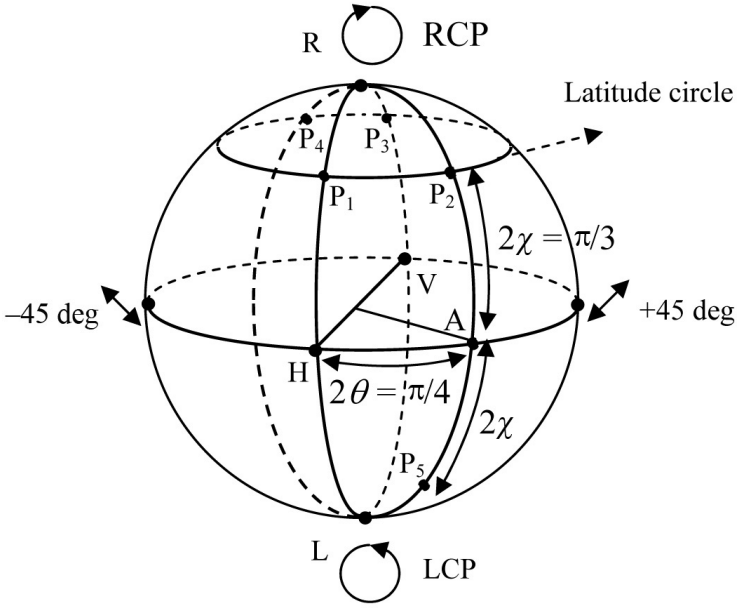


Figure 7.3(a) Various polarization states on the Poincaré sphere.

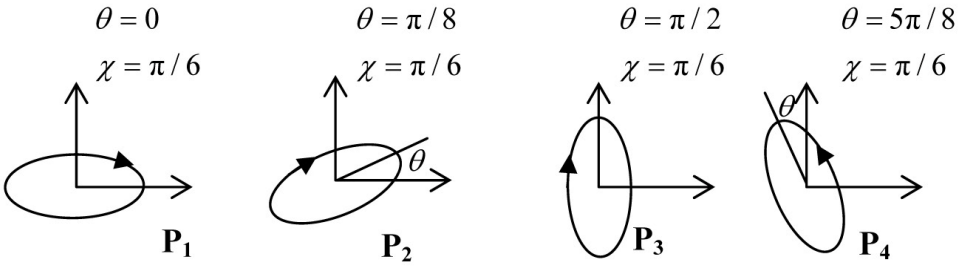


Figure 7.3(b) Polarization states represented by points P_1, P_2, P_3, and P_4 on the latitude circle, as shown in Fig. 7.3(a).

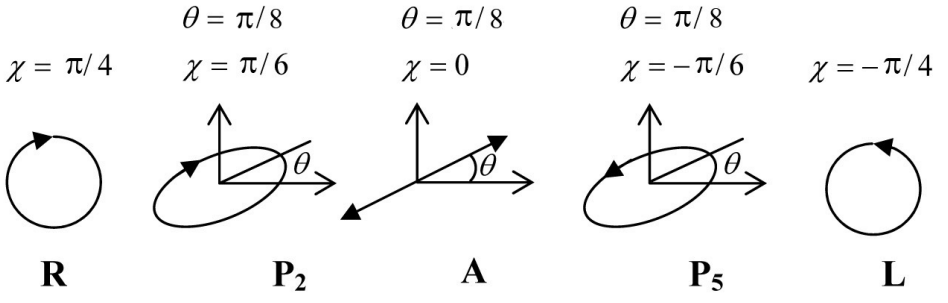


Figure 7.3(c) Polarization states represented by points R, P_2, A, P_5, and L on the meridian RAL, as shown in Fig. 7.3(a).

Example 7.1 Obtain the longitude and latitude of a right-elliptical SOP with $b/a = 1/\sqrt{3}$ and whose major axis makes an angle $\theta = \pi/4$ with the x axis.

Solution: The longitude will be $2\theta = \pi/2$. Since the sense of rotation is right, $\chi = \varepsilon = \tan^{-1}\left(\frac{b}{a}\right) = \tan^{-1}\left(\frac{1}{\sqrt{3}}\right) = \frac{\pi}{6}$. Thus, the latitude will be $2\chi = \frac{\pi}{3}$.

7.3.1 A polarizer and a birefringent medium

Polarizer: A polarizer is represented by a point on the Poincaré sphere that also represents the SOP produced by the polarizer. For example, in Fig. 7.2, point P also represents a polarizer producing an SOP characterized by orientation θ and ellipticity angle χ.

Birefringent medium: A birefringent medium is characterized by its two eigenpolarization states, known as *fast* and *slow* components, and is thus represented on the PS by the diameter FS, whose end points F and S represent the fast and slow SOP. For example, a linear birefringent medium such as a QWP whose fast and slow components are polarized along the horizontal and vertical directions, respectively, are represented by the diameter H–V [see Fig. 7.4(a)]. Similarly, an optically active medium in which fast and slow components are right- and left-circularly polarized are represented by the diameter R–L [see Fig. 7.4(b)].

7.4 Basic Properties of the Poincaré Sphere Representation

Using the PS representation, it is extremely simple to predict the effect of a polarizer/analyzer or a birefringent medium on a given SOP by using two basic

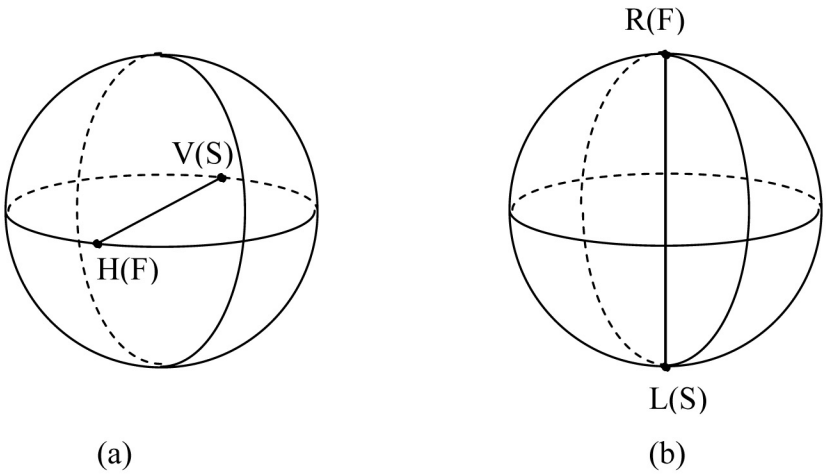


Figure 7.4 Representation of (a) a linear (with fast axis along the horizontal) and (b) a circular (with right-circular component as the fast component) birefringent medium on the Poincaré sphere.

properties of the PS representation, each of which is presented below with corresponding examples.

Property 1—effect of a polarizer/analyzer: If an SOP represented by a point P on a PS is passed through an analyzer represented by A, the fractional intensity transmitted by the analyzer is given by

$$I = \cos^2 \frac{\widehat{\mathrm{PA}}}{2}, \tag{7.2}$$

where $\widehat{\mathrm{PA}}$ is the arc length along the great circle connecting the two points on the Poincaré sphere (see Fig. 7.2). We verify this rule through the following simple examples.

Example 7.2 Obtain the fractional output intensity [Solution (i)] if a right- (or left-) circularly polarized beam is passed through a linear analyzer (LA), and [Solution (ii)] if a linearly polarized light is passed through a right- (or left-) circular polarizer.

Solution (i): A right- (or left-) circularly polarized state will be represented by the north (or south) pole on the PS, while the LA will be represented by some point A on the equator. It is clear that in this case $\widehat{\mathrm{PA}} = \pi/2$, which means that $I = \cos^2(\pi/4) = 1/2$, i.e., 50%.

Solution (ii): Similarly, if a linearly polarized light is passed through a right- (or left-) circular polarizer, again the fractional intensity passed through will be 50%.

Example 7.3 Obtain the output intensity if an elliptically polarized beam is passed through an LA whose pass axis is parallel to the major axis of the ellipse.

Solution: Let the elliptically polarized beam be represented by the point P (2θ, 2χ) on the PS (see Fig. 7.2). The LA with its pass axis parallel to the major axis of the ellipse will then be represented by the point A on the equator, lying on the meridian RPAL (Fig. 7.2). In this case, $\widehat{\mathrm{PA}} = 2\chi$, which means that the fractional intensity passed will be

$$I = \cos^2(\chi) = 1/(1+\tan^2\chi) = a^2/(a^2+b^2), \tag{7.3}$$

where we have used $\tan\chi = b/a$.

Property 2—effect of a birefringent medium: If an SOP given by P_i (θ_i, χ_i) is passed through a birefringent medium, the output SOP, P_o (θ_o, χ_o), can be obtained from the input state by rotating it clockwise on the sphere about the line FS representing the medium (clockwise while looking from F to S) by an angle δ,

where δ represents the phase difference introduced between the fast and the slow polarization states. This simple example demonstrates the use of the second property of the PS.

Example 7.4 Consider a linearly polarized wave making a 45-deg angle with the *x* axis. Obtain the output SOP if it is passed through a QWP with its fast axis parallel to the *x* axis.

Solution: On the PS, the given SOP will be represented by a point P and the QWP by the diameter HV, with H representing the fast axis, as shown in Fig 7.5, and $\overset{\frown}{\text{HP}} = \pi/2$. To obtain the output SOP, we rotate the sphere clockwise around diameter FS by an angle $\pi/2$, which will shift P to L. Thus, the output SOP will be left-circularly polarized.

In the following section, we show how simple it is to determine the effect of a birefringent medium on a given SOP using the PS representation. We consider the birefringent medium to be a QWP or an HWP and discuss its effect on various input states of polarization—namely, linear, circular, and elliptical.

7.4.1 Effect of a QWP/HWP on a linear SOP

Let the input SOP be polarized along the horizontal direction (*x* axis) so that it is represented by point H on the PS, and let the fast axis of the wave plate be oriented at an angle α with respect to the *x* axis (Fig. 7.6). As discussed earlier, the output SOP can be obtained by rotating the sphere around diameter FS

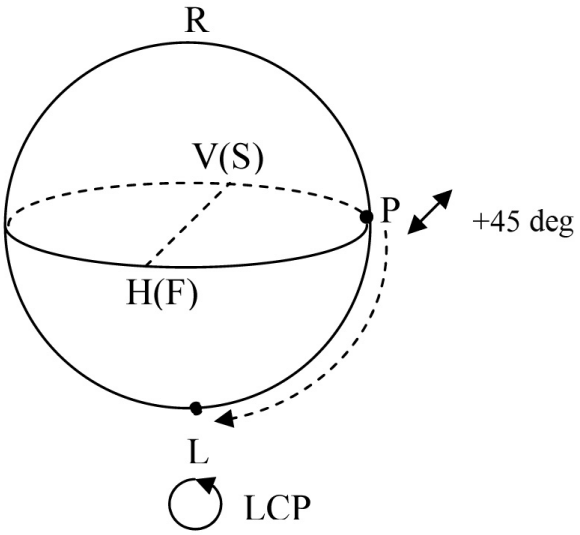


Figure 7.5 A QWP having the fast axis along the horizontal direction will convert a linear +45-deg SOP to a left-circular SOP.

clockwise by an angle of π/2 (phase difference introduced between the fast and slow polarizations by the wave plate). This means that the output SOP will be given by point Q (see Fig. 7.6) with coordinates (2α, 2α) representing a right-elliptically polarized light oriented at angle α with ellipticity b/a = tan (α). If $\alpha = \pi/4$, i.e., the fast axis of the QWP makes a 45-deg angle with the input SOP, the output SOP will be right-circularly polarized.

In the case of an HWP, the sphere will be rotated by an angle π. It is clear that the output SOP will be linear with longitude 4α on the PS, i.e., oriented at an angle 2α with the horizontal in physical space.

7.4.2 Effect of a QWP/HWP on a circular SOP

Let the input SOP be the right-circularly polarized wave represented by point R on the PS, and the fast axis of the wave plate be oriented at an angle α with respect to the *x* axis (Fig. 7.7). The output SOP can again be obtained by rotating the sphere clockwise around diameter FS by an angle of π/2 and will be given by point Q on the equator with longitude $2\alpha + \pi/2$ [see Fig. 7.7(a)]. Thus, the output SOP will be linearly polarized with orientation α + 45 deg. In a special case, if the fast axis is along the horizontal direction, the output SOP will be linearly polarized and oriented at 45 deg with respect to the horizontal. In the case of an HWP, the sphere will be rotated by an angle π, and the output SOP will be left-circularly polarized [see Fig. 7.7(b)].

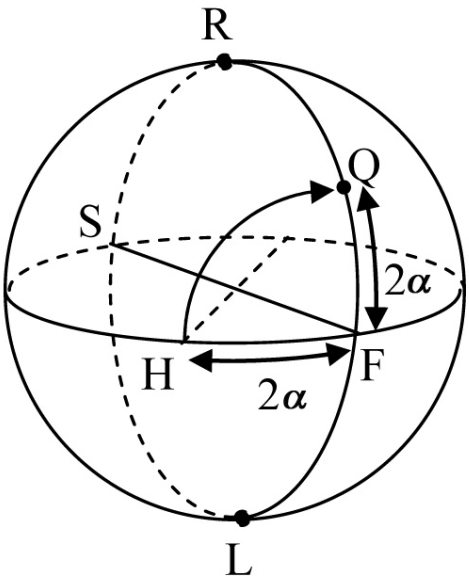


Figure 7.6 Effect of a QWP on linearly polarized light.

7.4.3 Effect of a QWP/HWP on an elliptical SOP

Let us now consider the general case of an elliptical SOP represented by a point P on the PS (Fig. 7.8). In this case, right-elliptically polarized light with orientation θ and ellipticity $\chi = \varepsilon$ is passed through a QWP represented by diameter FS. The output SOP corresponding to this input will, in general, be another elliptically polarized beam and can be obtained by rotating the sphere as discussed earlier.

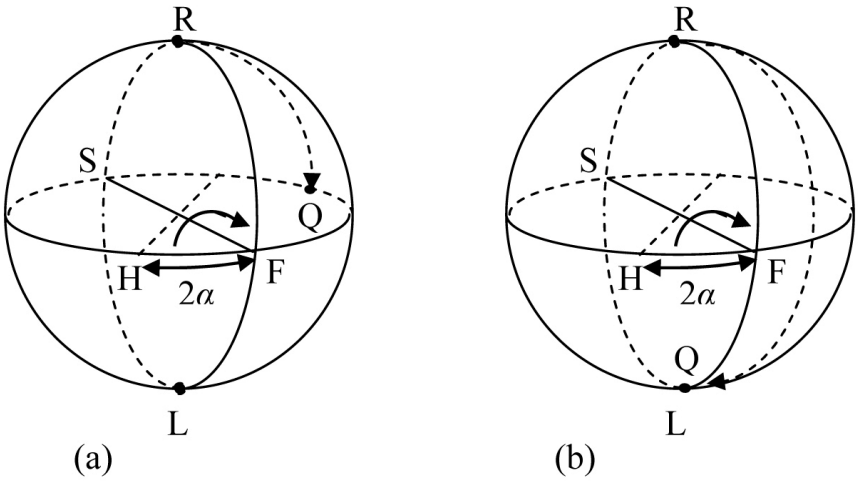


Figure 7.7 Effect of (a) a QWP and (b) an HWP on right-circularly polarized light.

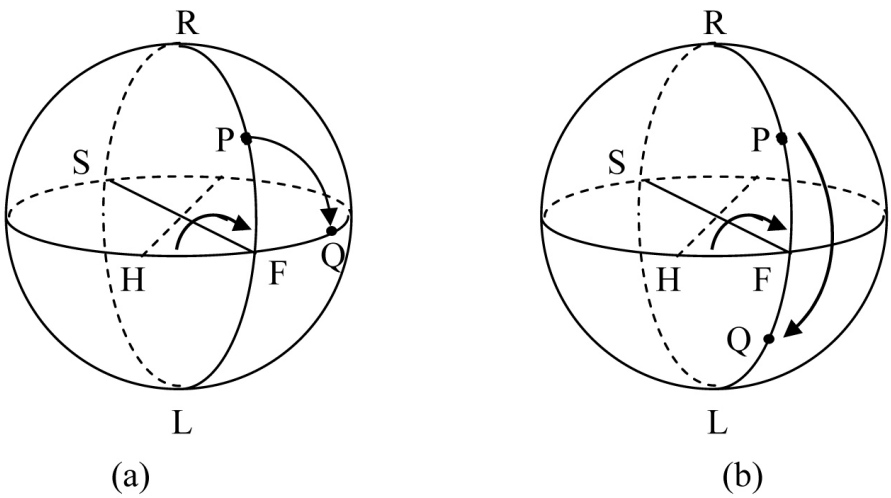


Figure 7.8 Effect of (a) a QWP and (b) an HWP on elliptically polarized light.

Here, we consider a specific case in which the fast axis of the wave plate is parallel to the major axis of the input ellipse. In this case, the end point F of the diameter FS representing the wave plate will coincide with the point where the meridian RPL passing through point P meets the equator, as shown in Fig. 7.8. The output SOP will then be a linear SOP represented by point Q on the equator such that FQ = 2ε, which will be oriented at an angle ε with respect to the fast axis of the wave plate. Thus, if right- (left-) elliptically polarized light of ellipticity (b/a) is passed through a QWP such that its fast axis coincides with the major axis of the ellipse, the output SOP will be linearly polarized and oriented at angle ε ($-\varepsilon$) with respect to the fast axis, where $\varepsilon = \tan^{-1}(b/a)$ (see Example 3.8).

In the case of an HWP, the sphere will be rotated by an angle π. It is clear that the output SOP will be elliptical, having the same orientation and ellipticity but the opposite sense of rotation.

The beauty of the PS representation lies in the fact that one can visualize a geometric picture of the evolution of the SOP through a birefringent medium, which makes understanding and designing the various polarization devices and experiments extremely easy. In order to demonstrate this, we discuss in the following two sections the working principles of some devices/experiments that use multiple optical components.

7.4.4 An ideal circular polarizer/analyzer

As described in Section 5.4.1, an *ideal circular polarizer* is a device that allows the desired circular component of the incident SOP to pass through it and stops the other component. Let us first consider the way in which an ideal right-circular polarizer can be realized. Using the PS, it is easy to understand that this can be achieved in the following way:

(i) Convert the right- and the left-circular components to linear components using a QWP; if the fast axis of the QWP is along the horizontal direction, the right- and left-circular components of the incident beam will be converted to linear +45-deg and linear −45-deg states, respectively [see Fig. 7.9(a)].

(ii) Stop the undesired component (linear −45 deg) by passing the beam through an LA with its pass axis along +45-deg direction.

(iii) Convert the desired component (linear +45 deg) back to the right-circular SOP using another QWP whose fast axis is along the vertical direction [see Fig. 7.9(b)].

7.4.5 An ideal elliptical polarizer/analyzer

An *elliptical polarizer/analyzer* is a device that allows a certain elliptical component of the incident SOP to pass through it and absorbs the orthogonal

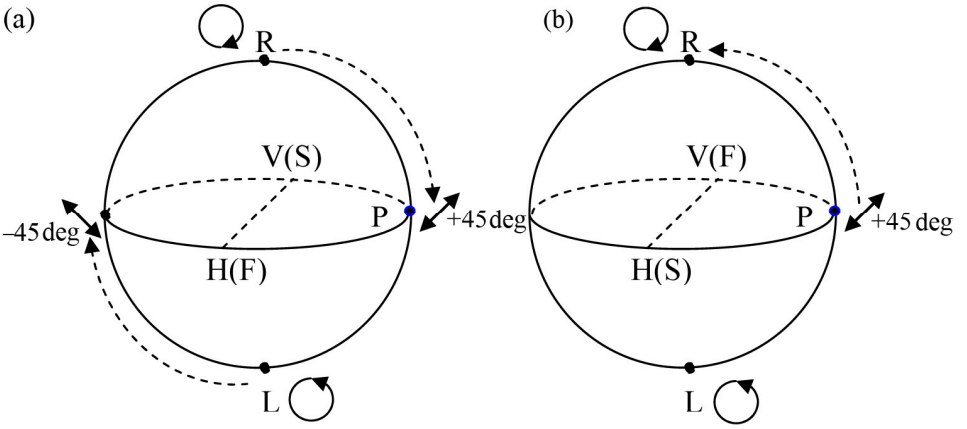


Figure 7.9 Action of the two QWPs in the realization of a right-circular polarizer/analyzer.

component. Let this desired component be represented by point P with parameters (2θ, 2χ) (see Fig. 7.10). We will show in the following that a combination of a QWP with its slow axis at θ, followed by an LA oriented at ($\theta - \chi$), and then followed by a QWP with its fast axis at an angle θ, will work as the desired ideal elliptical polarizer. This can be understood using the PS representation, as discussed below.

The elliptical component of a given SOP to be passed through the polarizer is represented by point P (2θ, 2χ) in Fig. 7.10(a). Let the meridian RPL meet the equator at point S (2θ, 0). It is clear from the figure that if the given SOP is passed through a QWP with its slow axis represented by S, its P component will be converted to a linear one oriented at ($\theta - \chi$) and represented by Q, which will completely pass through an LA oriented at ($\theta - \chi$). Finally, the second QWP with its fast axis at θ will convert this component again to the elliptical SOP represented by P, as shown in Fig. 7.10(b). The output SOP will thus be the required elliptical component of the input SOP. The component of the incident SOP that is orthogonal to P will be converted to a linear one orthogonal to the SOP represented by Q and will be eliminated by the LA represented by Q. Here we look at two more examples that show how the PS representation can be further used to determine SOPs.

Example 7.5 It is given that a light beam is elliptically polarized. How will you determine the complete SOP, i.e., orientation, ellipticity, and sense of rotation of the electric field vector with time? Explain your answer using PS representation.

Solution: Let the given SOP be represented by point P or Q on the PS, depending on whether the sense of rotation is right or left (see Fig. 7.11). We further assume that the orientation and the ellipticity of the given beam are given by θ_0 and b/a, respectively. Thus,

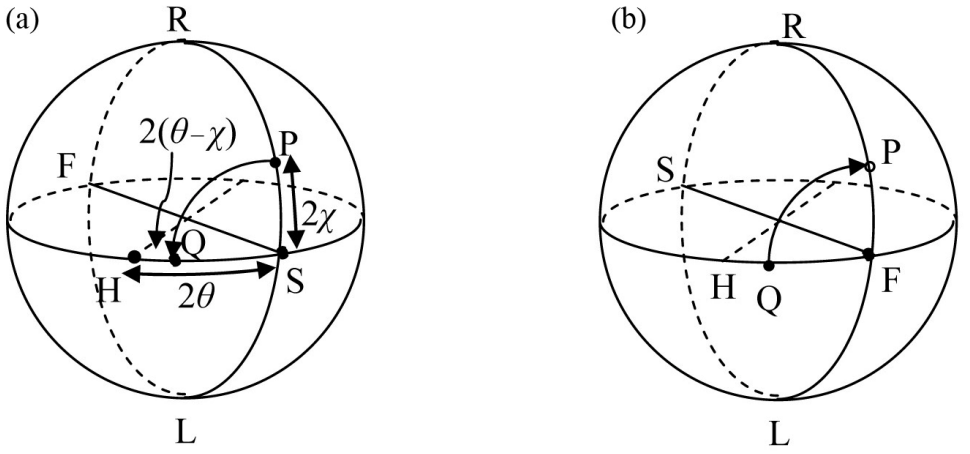


Figure 7.10 Function of the two QWPs in the realization of an elliptical polarizer/analyzer represented by P.

$$\overset{\frown}{\mathrm{HF}} = 2\theta_0 \text{ and } \overset{\frown}{\mathrm{PF}} = \overset{\frown}{\mathrm{FQ}} = 2\varepsilon,$$

where $\varepsilon = \tan^{-1}(b/a)$.

We first pass the given light beam through a rotating LA and note the output intensity as a function of orientation θ of the pass axis of the LA. On the PS, the analyzer will be represented by a point rotating on the equator. Obviously, the output intensity variation will be periodic in nature. Let $I_{\max}$ and $I_{\min}$ represent the maximum and minimum output intensities, which will correspond to the cases when the LA is represented by points F and S, respectively (see Fig. 7.11). It is easy to understand that the orientation θ of the analyzer in the case of maximum output intensity will be the same as that of the given SOP. Thus the orientation θ_0 is determined.

In order to determine the ellipticity and the sense of rotation of the given polarization state, we now insert a QWP in front of the rotating LA such that the fast axis of the QWP is oriented at θ_0. The output intensity is recorded again as a function of the orientation of the LA, which will again be periodic. However, now $I_{\min}$ will be 0, and $I_{\max}$ will appear at $\theta = \theta_0 + \varepsilon$ ($\theta_0 - \varepsilon$) if the sense of rotation is right (left), as discussed in the following, in which we use the PS representation to explain. It is clear from Fig. 7.11 that if the given SOP is passed through a QWP with its fast axis oriented at θ_0, the output SOP will become linear, represented by P_1 or Q_1 corresponding to the input P or Q such that

$$\overset{\frown}{\mathrm{FP}} = \overset{\frown}{\mathrm{FP_1}} = \overset{\frown}{\mathrm{FQ}} = \overset{\frown}{\mathrm{FQ_1}} = 2\varepsilon.$$

Thus, the output of the QWP when passed through a rotating LA will be periodic again, but with $I_{\min} = 0$ and the peak intensity shifted to $\theta \pm \varepsilon$. Hence, by

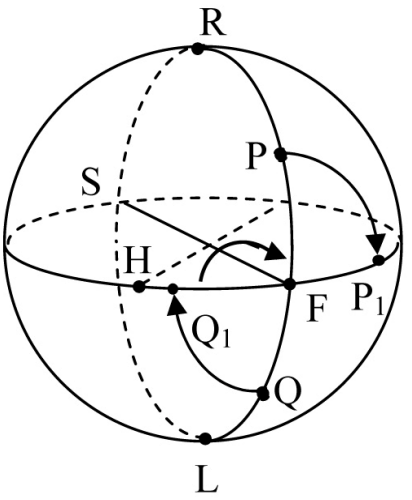


Figure 7.11 Effect of a QWP on an elliptic SOP when the fast axis of the QWP is parallel to the major axis of the ellipse.

measuring the shift in the peaks, the sense of rotation as well as the ellipticity can be determined.

Example 7.6 An unpolarized light beam is passed through a pair of crossed linear polarizers consisting of an HWP between them. If the wave plate rotates with an angular frequency ω, obtain the variation of the output intensity with time.

Solution: Let the crossed polarizers be represented by points H and V on the PS (Fig. 7.12). This means that the output SOP of the first polarizer will be represented by H and its intensity will be $I/2$, where I is the input intensity. Assuming that, at $t = 0$, if one of the axes (for instance, the slow axis) of the HWP is parallel to the pass axis along H, at time t, it will make an angle $\theta = \omega t$ with the horizontal direction, and hence on the PS, arc length $\overset{\frown}{\mathrm{HS}}$ will be $2\omega t$.

It is clear from the figure that at time t, the wave plate will convert the polarization state H to another linear state represented by Q such that arc length $\overset{\frown}{\mathrm{HQ}} = 4\omega$, and hence, arc length $\overset{\frown}{\mathrm{QV}}$ will be $\pi - 4\omega t$. Thus, the output intensity after passing through the second polarizer will be given by

$$I_o = \frac{I}{2}\cos^2\left(\frac{\pi - 4\omega t}{2}\right) = \frac{I}{4}[1 - \cos(4\omega t)] \ .$$

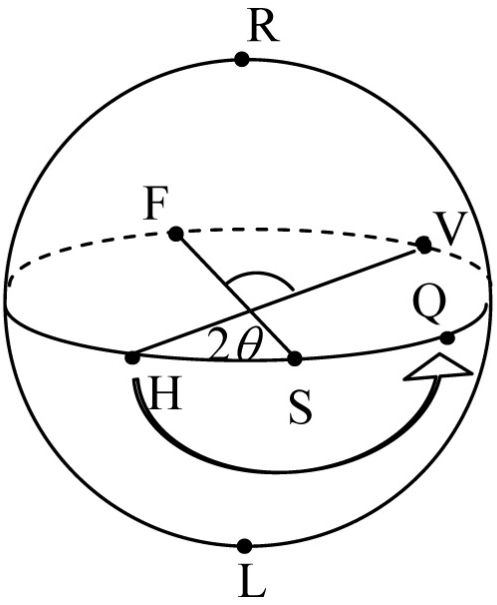


Figure 7.12 An HWP with its slow axis at S will convert the SOP from H to Q.

7.5 Poincaré Sphere and Stokes Parameters

Let us consider a polarized beam whose SOP is represented by a point P $(2\theta, 2\chi)$ on the PS. We now define a Cartesian coordinate system such that the center O represents the origin, and the *x*, *y*, and *z* axes are parallel to the lines joining O to the points representing the linear horizontal, linear +45-deg, and right-circular SOPs, respectively (see Fig. 7.13). The Cartesian coordinates of P will then be given by

$$\begin{aligned} x &= \cos 2\chi \cos 2\theta, \\ y &= \cos 2\chi \sin 2\theta, \\ z &= \sin 2\chi, \end{aligned} \tag{7.4}$$

which are simply the normalized Stokes parameters S_1, S_2, and S_3 of the given polarized light [see Eq. (6.33)] with S_0 = 1. Thus, the Stokes vector of a completely polarized beam can be thought of as a vector joining the center and the point on the PS.

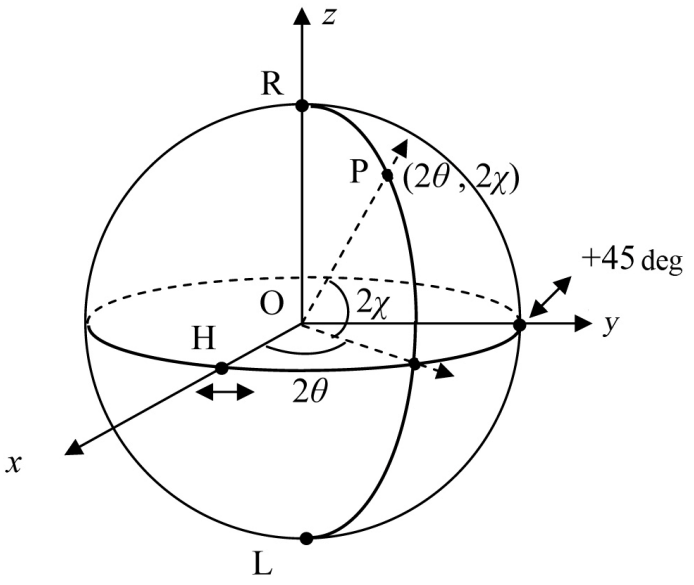


Figure 7.13 Cartesian coordinate system used to relate the PS representation and the Stokes parameters of a completely polarized wave.

Bibliography

Azzam, R. M. A. and N. M. Bashara, *Ellipsometry and Polarized Light*, North Holland Publishing Company, Amsterdam (1977).

Collett, E., *Field Guide to Polarization*, SPIE Press, Bellingham, WA (2005). [doi: 10.1117/3.626141].

Goldstein, D., *Polarized Light*, 2nd ed., Marcel Dekker, New York (2003).

Poincaré, H., Chapter 2 in *Théorie Mathématique de la Lumière*, vol. 2, Gauthiers-Villars, Paris (1892).

Ramachandran, G. N. and S. Ramaseshan, "Crystal optics," in *Encyclopedia of Physics*, **25/1**, S. Flugge, Ed., Springer, Berlin (1961).

Shurcliff, W. A., *Polarized Light: Production and Use*, Harvard University Press, Cambridge, MA (1962).

Theocaris, P. S. and E. E. Gdoutos, *Matrix Theory of Photoelasticity*, Springer Verlag, Berlin–Heidelberg (1979).

Chapter 8
Propagation and Polarization Characteristics of Single-Mode Fibers

8.1 Introduction

The invention of the laser in 1960 revolutionized the area of telecommunications, as it made it possible to modulate light waves at high frequencies. Initially, modulated laser beams propagating through open space were used to transmit information from one place to another. However, due to environmental disturbances, such as rain, fog, dust, temperature variations, and so on, such a transmission system was neither convenient nor reliable to use. Present-day optical communication systems use optical fibers through which information is transmitted in the form of optical pulses from one place to another. In the following, we discuss the basic propagation characteristics of optical fibers.

An *optical fiber* is a cylindrical structure, and the simplest (and the most extensively used) optical fiber is the *step-index fiber*, in which the refractive index distribution (in the transverse direction) is given by

$$n = \begin{cases} n_1 & 0 < r < a \\ n_2 & r > a, \end{cases} \tag{8.1}$$

where n_1 and n_2 ($< n_1$) represent the refractive indices of the core and the cladding, respectively, and a represents the radius of the core (see Fig. 8.1).

The cladding is usually pure silica, while the core is usually silica doped with germanium; doping by germanium results in an increase in refractive index. The guidance of the light beam (through the optical fiber) takes place because of the total internal reflection at the core–cladding interface. The rays for which the angle ϕ is more than the critical angle ϕ_c, where

$$\phi_c = \sin^{-1}\left(\frac{n_2}{n_1}\right) \tag{8.2}$$

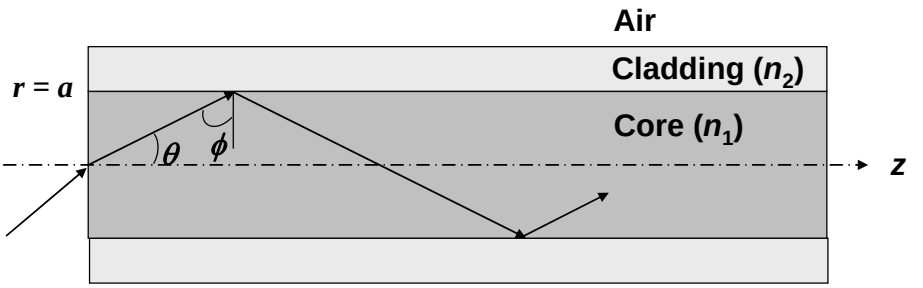


Figure 8.1 A step-index fiber. Light guidance takes place through the phenomenon of total internal reflection at the core–cladding interface.

will be totally internally reflected and will be guided through the core. The preceding condition of guidance can also be put in terms of angle θ made by the ray with the fiber axis. For guided rays,

$$\theta \leq \cos^{-1}\left(\frac{n_2}{n_1}\right). \tag{8.3}$$

When a light pulse propagates through an optical fiber, it suffers from attenuation due to various mechanisms, and the pulse broadens in time, leading to what is termed *pulse*/*chromatic dispersion*. Attenuation and chromatic dispersion represent the two most important characteristics that determine the information transmission capacity of optical fibers. Obviously, the lower the attenuation (and similarly, the lower the dispersion), the greater will be the required repeater spacing and therefore the higher the information-carrying capacity and the lower the cost of the communication system.

8.2 Attenuation in Optical Fibers

The attenuation in an optical fiber is specified in terms of the unit dB/km and is defined as

$$\alpha\ (\text{dB/km}) = \frac{10}{L\ (\text{km})} \log_{10}\left(\frac{P_{\text{in}}}{P_{\text{out}}}\right). \tag{8.4}$$

Here P_{in} and P_{out} are the input and output powers corresponding to an optical fiber of length L (km). The losses are caused by various mechanisms, such as Rayleigh scattering, absorption due to metallic impurities, and the presence of a small amount of water, as well as due to the intrinsic absorption of silica molecules. Even 1 part per million (ppm) of iron can cause a loss of about 0.68 dB/km at 1100 nm. Similarly, a concentration of 1 ppm of OH^- ion can cause a loss of 4 dB/km at 1380 nm. This shows the level of purity that is required to

achieve very low-loss optical fibers. The peaks shown in Fig. 8.2 are due to a very minute presence of OH^- ions as well as to metal impurities.

The lowest loss (~ 0.25 dB/km) occurs at about 1550 nm (the approximate wavelength[1] of erbium-doped fiber amplifiers); therefore, most optical communication systems today operate at this wavelength.

8.3 Modes of a Step-Index Fiber

In Section 8.1 we used ray optics to study propagation of a lightwave through a step-index fiber; however, this description is only approximate and cannot be used to understand the polarization characteristics of an optical fiber. The exact approach is to obtain the solutions of Maxwell's equations that satisfy the proper boundary conditions at the core–cladding interface, which are known as the *modes* of the optical fiber. The boundary conditions are the continuity of (i) the tangential components of **E** and **H**, and (ii) the normal components of **B** and **D** vectors at $r = a$. Physically, the modes are various transverse field distributions that propagate along the fiber without any change in their transverse field distributions.

The exact vector modes of optical fibers are a bit complicated, and their mathematical description is difficult. A system engineer who has not gone through a rigorous course on the theory of optical fibers will find it difficult to

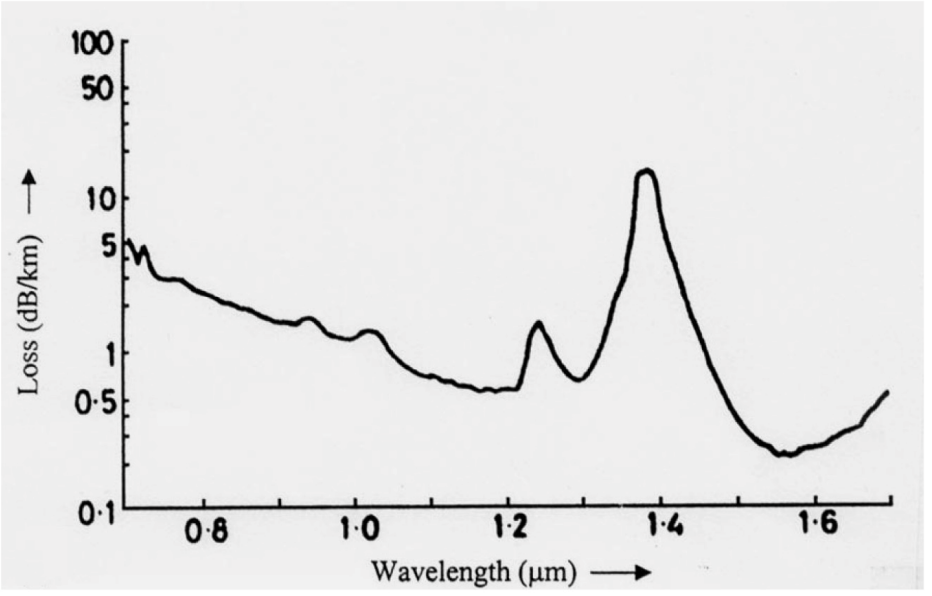


Figure 8.2 Typical wavelength dependence of loss for a silica fiber. The peaks in the attenuation curve in the wavelength regions 1.25 and 1.40 μm are due to the presence of minute amounts of water and other impurities. Notice that the lowest loss occurs at 1550 nm. (Reprinted with permission from Ref. 1, © 1979, IEEE.)

understand these modes. Fortunately, for practical single-mode fibers, the core and cladding refractive index difference $n_1 - n_2$ is extremely small, and one can use the weakly guiding approximation ($n_1 \approx n_2$) to obtain the propagation constants of the various modes, which makes the calculations extremely simple.

This approximation is also expressed in terms of a parameter Δ, defined as

$$\Delta \equiv \frac{n_1^2 - n_2^2}{2n_1^2}. \tag{8.5}$$

For $n_1 \approx n_2$, $\Delta << 1$, and

$$\Delta = \frac{n_1 - n_2}{n_1}\frac{n_1 + n_2}{2n_1} \approx \frac{n_1 - n_2}{n_2} \approx \frac{n_1 - n_2}{n_1}. \tag{8.6}$$

The error introduced by the weakly guiding approximation in the calculation of fiber modes is so small that in practice, the exact modes are rarely used. In this approximation, the modes of a fiber are linearly polarized and are known as linearly polarized (LP) modes, which are discussed in the following.

8.3.1 Linearly polarized (LP) modes

As discussed in Chapter 2, the solutions to Maxwell's equations in an infinitely extended homogeneous medium are transverse in nature and linearly polarized. In the weakly guiding approximation, $n_1 \approx n_2$, we assume that the modes of an optical fiber maintain the previously mentioned nature. Accordingly, we assume that the transverse component of the electric field (E_x or E_y) satisfies the following scalar wave equation:

$$\nabla^2 \Psi = \frac{n^2}{c^2}\frac{\partial^2 \Psi}{\partial t^2}, \tag{8.7}$$

where Ψ represents E_x (or E_y). Further, instead of satisfying the exact boundary conditions (continuity of E_z, $n^2 E_r$, and E_ϕ at $r = a$), we satisfy the continuity of E_x (or E_y) and $\partial E_x/\partial r$ (or $\partial E_y/\partial r$ at $r = a$. We solve Eq. (8.7) in the cylindrical system of coordinates and write the solution of Eq. (8.7) in the form

$$\Psi(r, \phi, z, t) = \psi(r, \phi) e^{i(\omega t - \beta z)}, \tag{8.8}$$

where β is the propagation constant.

Substituting Ψ from Eq. (8.8) into Eq. (8.7), one can easily show that $\psi(r,\phi)$ satisfies the following equation:

$$\frac{\partial^2\psi}{\partial r^2}+\frac{1}{r}\frac{\partial\psi}{\partial r}+\frac{1}{r^2}\frac{\partial^2\psi}{\partial\phi^2}+\left[k_0^2n^2(r)-\beta^2\right]\psi=0. \tag{8.9}$$

Equation (8.9) can be solved by using the method of separation of variables, i.e., $\psi(r,\phi)$ can be assumed to be of the form

$$\psi(r,\phi) = R(r)\Phi(\phi). \tag{8.10}$$

Substituting $\psi(r,\phi)$ from Eq. (8.10) into Eq. (8.9), one obtains

$$\frac{r^2}{R}\left(\frac{d^2R}{dr^2}+\frac{1}{r}\frac{dR}{dr}\right)+r^2\left[k_0^2n^2(r)-\beta^2\right]=-\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\phi^2} = l^2, \tag{8.11}$$

where l is a constant. Since the medium has cylindrical symmetry, the ϕ dependence of $\psi(r,\phi)$ should be of the form $\cos l\phi$ or $\sin l\phi$, and for the function to be single-valued, i.e., $\Phi(\phi+2\pi)=\Phi(\phi)$, we must have l = 0, 1, 2, …. Thus, $\psi(r,\phi)$ is given by

$$\psi(r,\phi) = R(r)\begin{Bmatrix}\cos l\phi\\ \sin l\phi\end{Bmatrix}, \tag{8.12}$$

where $R(r)$ satisfies the radial part of Eq. (8.9):

$$r^2\frac{d^2R}{dr^2}+r\frac{dR}{dr}+\left\{\left[k_0^2n^2(r)-\beta^2\right]\ r^2-l^2\right\}R=0. \tag{8.13}$$

Before we solve Eq. (8.13), we will make some general comments about its solutions, which can be divided into two distinct classes:

(i) *Guided modes*: For guided modes,

$$k_0n_2<\beta<k_0n_1. \tag{8.14}$$

For β lying in the preceding range, the field $R(r)$ is oscillatory in the core and decays in the cladding, and β assumes only discrete values; these are known as the *guided modes* of the waveguide. For a given value of l, there will be several guided modes.

(ii) *Radiation modes*: For radiation modes, $\beta < k_0 n_2$, the fields are oscillatory in the core as well as cladding, and β can assume a continuum of values. These are known as the *radiation modes.* Such modes leak away quickly and hence are not important in a long-distance communication system. In the analysis provided in this section, we will concentrate only on the guided modes.

For a step-index fiber, the well-behaved solutions of Eq. (8.13), which are continuous at the $r = a$ interface, can be written in terms of Bessel functions[2–4] J_l and K_l as

$$R(r) = \begin{cases} \dfrac{A}{J_l(U)} J_l\left(\dfrac{Ur}{a}\right) & r < a \\ \dfrac{A}{K_l(W)} K_l\left(\dfrac{Wr}{a}\right) & r > a. \end{cases} \tag{8.15}$$

Thus,

$$\psi(r,\phi) = \begin{cases} \dfrac{A}{J_l(U)} J_l\left(\dfrac{Ur}{a}\right)\begin{bmatrix}\cos l\phi \\ \sin l\phi\end{bmatrix} & r < a \\ \dfrac{A}{K_l(W)} K_l\left(\dfrac{Wr}{a}\right)\begin{bmatrix}\cos l\phi \\ \sin l\phi\end{bmatrix} & r > a, \end{cases} \tag{8.16}$$

where A is a constant, and we have assumed the continuity of ψ at the core–cladding interface. The constants U and W are defined as

$$U = a\sqrt{k_0^2 n_1^2 - \beta^2} \text{ and } W = a\sqrt{\beta^2 - k_0^2 n_2^2}\ . \tag{8.17}$$

For guided modes, $n_2^2 k_0^2 < \beta^2 < n_1^2 k_0^2$, and therefore, both U and W are real.

The normalized waveguide parameter V is defined as

$$V = \sqrt{U^2 + W^2} = k_0 a\sqrt{n_1^2 - n_2^2}\ . \tag{8.18}$$

The waveguide parameter V (which also depends on the operating wavelength λ_0) is an extremely important quantity characterizing an optical fiber.

If we make $d\psi / dr$ continuous at $r = a$, we obtain the eigenvalue equation for LP modes as

$$\frac{UJ_l'(U)}{J_l(U)} = \frac{WK_l'(W)}{K_l(W)}, \tag{8.19}$$

where $J_l'(U)$ and $K_l'(W)$ represent the derivatives of $J_l(U)$ and $K_l(W)$ with respect to their arguments. Using the following recurrence relations of Bessel functions:

$$J_l'(x) = \mp \frac{lJ_l(x)}{x} \pm J_{l\mp1}(x) \tag{8.20a}$$

$$K_l'(x) = \mp \frac{lK_l(x)}{x} - K_{l\mp1}(x), \tag{8.20b}$$

Eq. (8.19) can be recast as

$$\frac{UJ_{l-1}(U)}{J_l(U)} = -\frac{WK_{l-1}(W)}{K_l(W)}. \tag{8.21}$$

It is convenient to define the normalized propagation constant as

$$b = \frac{\frac{\beta^2}{k_0^2} - n_2^2}{n_1^2 - n_2^2} = \frac{W^2}{V^2}. \tag{8.22}$$

Thus, $W = V\sqrt{b}$, and $U = V\sqrt{1-b}$. In terms of b, the eigenvalue Eq. (8.21) can be written as[2]

$$V\sqrt{1-b}\,\frac{J_{l-1}\left[V\sqrt{1-b}\right]}{J_l\left[V\sqrt{1-b}\right]} = -V\sqrt{b}\,\frac{K_{l-1}\left[V\sqrt{b}\right]}{K_l\left[V\sqrt{b}\right]} \qquad l \geq 1, \tag{8.23}$$

and

$$V\sqrt{1-b}\,\frac{J_1\left[V\sqrt{1-b}\right]}{J_0\left[V\sqrt{1-b}\right]} = V\sqrt{b}\,\frac{K_1\left[V\sqrt{b}\right]}{K_0\left[V\sqrt{b}\right]} \qquad l = 0. \tag{8.24}$$

For a given value of V, the solution of the preceding transcendental equations will give us discrete values of the normalized propagation constant b of the guided LP_{lm} modes. For a given value of l, there will be finite number of

solutions, and the m'th solution (m = 1, 2, 3, …) is referred to as the LP_{lm} mode. An interesting property of LP_{lm} modes is that one obtains universal curves describing the dependence of b (and therefore of U and W) on V. Further, since the modes are solutions of the scalar wave equation, they satisfy the orthonormality condition

$$\int_0^{\infty}\int_0^{2\pi} \psi^{*}_{lm}(r,\phi)\,\psi_{l'm'}(r,\phi)\,r\;dr\;d\phi = \delta_{ll'}\delta_{mm'}, \tag{8.25}$$

where $\delta_{mm'}$ is the Kronecker delta function.

8.3.2 Cutoff V values of LP_{lm} modes

It should be noted that a particular mode is possible only if the V value of the fiber is greater than a certain value, say, V_c, which is known as the *cutoff V value* of that mode. Physically, at the cutoff of a mode, the angle of incidence of the corresponding ray at the core–cladding interface is equal to the critical angle; hence,

$$\beta = k_0 n_2 \text{, i.e., } W = 0 \text{, giving } U = V_c. \tag{8.26}$$

Thus, the cutoff V values of LP_{lm} modes can be obtained by substituting $\beta = k_0 n_2$ [i.e., $W = 0$ and $U = V_c$ in Eqs. (8.23) and (8.24)], which is the m'th root of

(i) $$J_1(V_c) = 0\,,\; l = 0 \text{ modes,} \tag{8.27a}$$

(ii) $$J_0(V_c) = 0\,,\; l = 1 \text{ modes,} \tag{8.27b}$$

(iii) $$J_{l-1}(V_c) = 0\,,\; V_c \neq 0\,,\; l \geq 2 \text{ modes.} \tag{8.27c}$$

Using Eqs. (8.27a)–(8.27c), the cutoff V values of the first few LP_{lm} modes are given here: $LP_{01} \to 0.00$; $LP_{11} \to 2.4048$; LP_{02}, $LP_{21} \to 3.8317$.

Figure 8.3 gives the universal curves showing the variation of b with V for the few lower-order modes obtained from Eqs. (8.23) and (8.24). Using Eqs. (8.14) and (8.22), we see that guided modes are characterized by $0 < b < 1$. As can be seen, at a particular V value, the fiber can support only a finite number of modes.

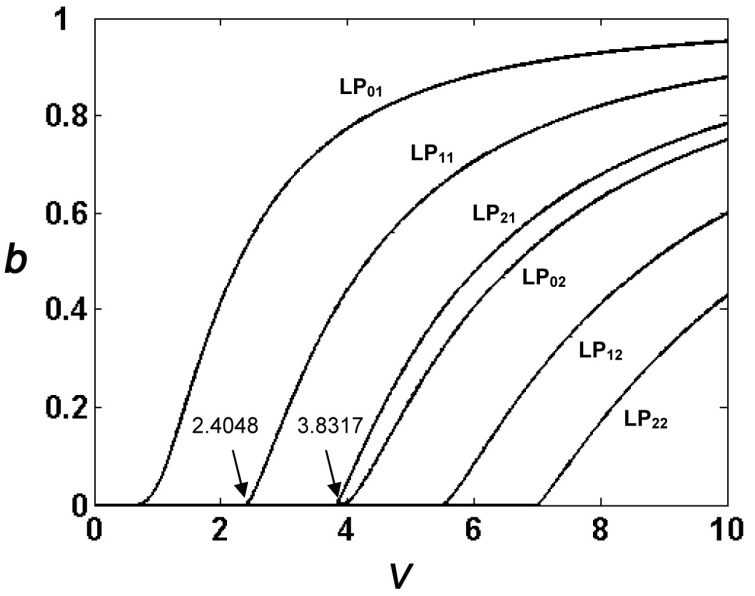


Figure 8.3 Variation of the normalized propagation constant *b* with normalized waveguide parameter *V* corresponding to a few lower-order modes. (Calculations courtesy of Ms. Triranjita Srivastava.)

8.4 Single-Mode Fiber

As is evident from Fig. 8.3, for a step-index fiber with $0 < V < 2.4048$, we will have only one guided mode—namely, the LP_{01} mode, also referred to as the *fundamental mode*. Such a fiber is referred to as a *single-mode fiber* (SMF) and is of tremendous importance in optical fiber communication systems. Polarization characteristics of optical fibers are also important, mainly only in the single-mode region. Since V depends on the operating wavelength, a given fiber may be single mode at one wavelength, while it may support more than one mode at another wavelength. The wavelength for which $V = 2.4048$ is known as the *cutoff wavelength* and is denoted by λ_c. Thus, a given fiber is single mode when $\lambda_0 > \lambda_c$. For a typical single-mode fiber (operating at 1550 nm), $a = 4.2$ μm, $n_2 = 1.444$ (for pure silica), and $\Delta = 0.0034$.

Example 8.1 As an example, we consider a step-index fiber with

$$n_2 = 1.447,\ \Delta = 0.003, \text{ and } a = 4.2\ \mu\text{m},$$

giving

$$V \simeq k_0 a n_2 \sqrt{2\Delta} = \frac{2.958}{\lambda_0},$$

where λ_0 is measured in μm. In this example, λ_c is given by

$$\frac{2.958}{\lambda_c} = 2.4048 \Rightarrow \quad \lambda_c = 1.23 \ \ \mu\text{m}.$$

Thus, for $\lambda_0 > 1.23$ μm, the fiber will be single mode.

8.4.1 Modal field pattern of the fundamental mode

The transverse dependence of the modal field of the fundamental mode is given by

$$\psi(r,\phi) = \begin{cases} \dfrac{A}{J_0(U)} J_0\left(\dfrac{Ur}{a}\right) & r < a \\ \dfrac{A}{K_0(W)} K_0\left(\dfrac{Wr}{a}\right) & r > a, \end{cases} \tag{8.28}$$

where U and W are determined by solving Eq. (8.24). For a single-mode step-index fiber, a convenient empirical formula for $b(V)$ is given by[5]

$$b(V) = \left(A - \frac{B}{V}\right)^2 \quad 1.5 \lesssim V \lesssim 2.5, \tag{8.29}$$

where $A \approx 1.1428$ and $B \approx 0.996$.

Table 8.1 gives the values of b as a function of V for the fundamental (LP_{01}) mode as obtained from Eqs. (8.24) and (8.29), respectively, for a step-index fiber.

Table 8.1 Values of b versus V of the LP_{01} mode for a step-index fiber.

V	b [by solving Eq. (8.24)]	b [using Eq. (8.29)]
1.5	0.229248	0.229249
1.6	0.270063	0.270712
1.7	0.309467	0.310157
1.8	0.347068	0.347471
1.9	0.382660	0.382653
2.0	0.416163	0.415767
2.1	0.447581	0.446911
2.2	0.476969	0.476200
2.3	0.504416	0.503754
2.4	0.530026	0.529693
2.5	0.553915	0.554131

Once b is known, one can calculate the propagation constant β using Eq. (8.22). Now, from Eq. (8.22),

$$b = \frac{\left(\frac{\beta}{k_0} - n_2\right)\left(\frac{\beta}{k_0} + n_2\right)}{(n_1 - n_2)(n_1 + n_2)}. \tag{8.30}$$

Since for a guided mode, β/k_0 lies between n_1 and n_2, and for all practical single-mode fibers, n_1 is very close to n_2, we may write Eq. (8.30) as

$$b \approx \frac{\frac{\beta}{k_0} - n_2}{n_1 - n_2}, \tag{8.31}$$

giving

$$\beta = \frac{\omega}{c}\left[n_2 + \left(n_1 - n_2\right) b\left(V\right)\right] = \frac{\omega}{c} n_2 (1 + \Delta b), \tag{8.32}$$

where we have used $\Delta = (n_1 - n_2) / n_2$.

Example 8.2 We consider a step-index fiber with $n_2 = 1.447$, $\Delta = 0.003$, and $a = 4.2$ μm, giving $V = 2.958/\lambda_0$, where λ_0 is measured in μm. Thus, for $\lambda_0 > 1.23$ μm, the fiber will be single mode. We assume an operating wavelength $\lambda_0 = 1.479$ μm so that $V = 2.0$, and therefore (from Table 8.1),

$$b \approx 0.4162 \;\Rightarrow\; \frac{\beta}{k_0} \approx n_2\,(1 + \Delta b) \approx 1.4488$$

$$\Rightarrow\; \beta \approx 6.1549 \times 10^6 \text{ m}^{-1}.$$

8.4.2 Spot size of the fundamental mode

The transverse field distribution associated with the fundamental mode of a single-mode fiber is an extremely important quantity, and it determines various important parameters, such as splice loss at joints between fibers, launching efficiencies from sources, and bending loss. For a step-index fiber, one obtains analytical expressions for the fundamental-mode field distribution in terms of Bessel functions [Eq. (8.28)]. For most single-mode fibers with a general transverse refractive index profile, the fundamental-mode field distribution can be well approximated by a Gaussian function, which can be written in the form

$$\psi(x, y) = A e^{-\frac{x^2+y^2}{w^2}} = A e^{-\frac{r^2}{w^2}}, \tag{8.33}$$

where w is referred to as the *Gaussian spot size* of the mode field pattern and $2w$ is called the *Gaussian mode field diameter* (MFD). MFD is a very important characteristic of a single-mode optical fiber. For a step-index fiber, one often uses the following empirical expression[6] for w:

$$\frac{w}{a} \approx 0.65 + \frac{1.619}{V^{3/2}} + \frac{2.879}{V^6} \qquad 0.8 \leq V \leq 2.5, \tag{8.34}$$

where a is the core radius. As an example, for the step-index fiber considered in the preceding examples and operating at 1300 nm, we have $V \approx 2.28$, giving $w \approx 4.8$ μm. Note that the spot size is larger than the core radius of the fiber; this is due to the penetration of the modal field into the cladding of the fiber. The same fiber will have a V value of 1.908 at $\lambda_0 = 1550$ nm, giving a value of the spot size ≈ 5.5 μm. *Thus, in general, the spot size increases with wavelength.* The standard single-mode fiber (SSMF) designated as G-652 fiber for operation at 1310 nm has an MFD of 9.2 ± 0.4 μm and an MFD of 10.4 ± 0.8 μm at 1550 nm.

For $2.4048 < V < 3.8317$, a step-index fiber supports two modes—namely, LP_{01} and LP_{11}—and the fiber is known as a two-mode fiber. For larger values of V (≥10), the fiber supports many modes and is known as a multimode fiber. Different modes (in a multimode fiber) travel with different group velocities, creating what is known as *intermodal dispersion*.

8.5 Pulse Dispersion in Single-Mode Optical Fibers

In digital communication systems, information to be sent is first coded in the form of pulses of light, and then these pulses are transmitted from the transmitter to the receiver, where the information is decoded. The larger the number of pulses that can be sent per unit time and remain resolvable at the receiver end, the larger the transmission capacity of the system. A pulse of light sent into a fiber broadens in time as it propagates through the fiber; this phenomenon is known as *pulse dispersion* (see Fig. 8.4). Pulse dispersion can occur for a variety of reasons. In the next section we will discuss the broadening of optical pulses due to the wavelength dependence of the refractive index of the fiber material, known as *material dispersion*.

8.5.1 Material dispersion

If β represents the propagation constant of the mode, the pulse will propagate through the fiber with what is known as the *group velocity* v_g, which is given by[7]

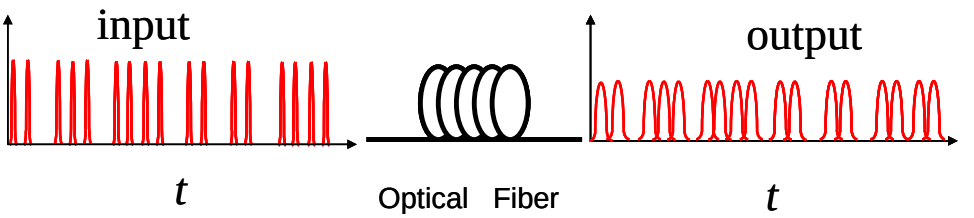


Figure 8.4 Schematic drawing of pulse dispersion: coded information in the form of light pulses is propagated through the optical fiber, where the pulses are attenuated and broadened in time.

$$v_g = \frac{1}{d\beta / d\omega}. \tag{8.35}$$

As most of the power carried by the mode is confined in the core, we assume that the pulse propagating through the fiber sees mainly the core of refractive index $n_1(\omega)$, and hence,

$$\beta(\omega) = \frac{\omega}{c} n_1(\omega). \tag{8.36}$$

Thus,

$$\frac{1}{v_g} = \frac{d\beta}{d\omega} = \frac{1}{c}\left[n_1(\omega) + \omega \frac{dn_1}{d\omega} \right]. \tag{8.37}$$

We may mention that it is customary to express $n(\omega)$ in terms of the free-space wavelength λ_0, which is related to ω through the equation

$$\omega = \frac{2\pi c}{\lambda_0}. \tag{8.38}$$

Thus,

$$\frac{dn_1}{d\omega} = \frac{dn_1}{d\lambda_0}\frac{d\lambda_0}{d\omega} = -\frac{\lambda_0^2}{2\pi c}\frac{dn_1}{d\lambda_0},$$

or

$$\frac{1}{v_g} = \frac{1}{c}\left[n_1(\lambda_0) - \lambda_0 \frac{dn_1}{d\lambda_0} \right]. \tag{8.39}$$

Thus, the time taken by a pulse to traverse a fiber length L is given by

$$\tau = \frac{L}{v_g} = \frac{L}{c}\left[n_1(\lambda_0) - \lambda_0 \frac{dn_1}{d\lambda_0} \right]. \tag{8.40}$$

Since its right-hand side depends on λ_0, Eq. (8.40) implies that different wavelengths will travel with different group velocities while propagating through a certain length of the fiber. Every source of light has a certain wavelength spread, which is often referred to as the *spectral width of the source*. For example, a white-light source (such as the sun) has a spectral width of about 300 nm; whereas, a light-emitting diode (LED) has a spectral width of about 25 nm, and a typical laser diode (LD) operating at 1300 nm has a spectral width of about 2 nm or less. Each wavelength component of the pulse travels with a slightly different group velocity and, in general, results in the broadening of the pulse. Thus, the pulse broadening is given by

$$\Delta\tau_m = \frac{d\tau}{d\lambda_0}\Delta\lambda_0,$$

or

$$\Delta\tau_m = -\frac{L\,\Delta\lambda_0}{c\,\lambda_0}\left[\lambda_0^2 \frac{d^2 n_1}{d\lambda_0^2} \right], \tag{8.41}$$

where L is the length of the fiber, $\Delta\lambda_0$ is the spectral width of the source, and c is the speed of light in free space. The subscript m in Eq. (8.41) signifies that we are considering material dispersion. We assume that

$$\Delta\lambda_0 = 1 \text{ nm} = 10^{-9} \text{ m} \text{ and } L = 1 \text{ km} = 1000 \text{ m},$$

and define the dispersion coefficient as

$$D_m = \frac{\Delta\tau_m}{L\Delta\lambda_0} \approx -\frac{1}{3\lambda_0}\left(\lambda_0^2 \frac{d^2 n_1}{d\lambda_0^2} \right) \times 10^4 \quad \text{ps/(nm} \cdot \text{km)}, \tag{8.42}$$

where λ_0 is measured in μm and we have assumed that c ≈ 3 × 10^8 m/s. The quantity D_m is usually referred to as the *material dispersion coefficient*. A medium is said to be characterized by positive dispersion when D_m is positive, and is said to be characterized by negative dispersion when D_m is negative.

For fused silica, which is the basic material used in optical fibers, an accurate empirical relation for refractive-index variation with wavelength, known as the *Sellmeier equation*, is given by[8]

$$n(\lambda_0) = C_0 + C_1\lambda_0^2 + C_2\lambda_0^4 + \frac{C_3}{(\lambda_0^2 - l)} + \frac{C_4}{(\lambda_0^2 - l)^2} + \frac{C_5}{(\lambda_0^2 - l)^3}, \quad (8.43)$$

where $C_0 = 1.4508554$, $C_1 = -0.0031268$, $C_2 = -0.0000381$, $C_3 = 0.0030270$, $C_4 = -0.0000779$, $C_5 = 0.0000018$, $l = 0.035$, and λ_0 is measured in μm. Using Eqs. (8.42) and (8.43), one can obtain the contribution of material dispersion in an optical fiber.

Example 8.3 In the first-generation optical communication systems, one used LEDs with $\lambda_0 \approx 0.85$ μm and $\Delta\lambda_0 \approx 25$ nm. At $\lambda_0 \approx 0.85$ μm, for pure silica,

$$\frac{d^2n}{d\lambda_0^2} \approx 0.030 \ (\mu\text{m})^{-2},$$

giving

$$D_m \approx -85 \text{ ps/(nm·km)},$$

the negative sign indicating that higher wavelengths travel faster than lower wavelengths. Thus, for $\Delta\lambda_0 \approx 25$ nm, the actual broadening of the pulse will be

$$\Delta\tau_m \approx 2.1 \text{ ns/km},$$

implying that, because of material dispersion, the pulse will broaden by 2.1 ns after traversing through 1 km of silica fiber.

Example 8.4 In the fourth-generation optical communication systems, one uses laser diodes at $\lambda_0 \approx 1.55$ μm with $\Delta\lambda_0 \approx 2$ nm. Now, at $\lambda_0 \approx 1.55$ μm,

$$\frac{d^2n}{d\lambda_0^2} \approx -0.0042 \ (\mu\text{m})^{-2},$$

giving

$$D_m \approx +21.7 \text{ ps/(nm·km)},$$

the positive sign indicating that higher wavelengths travel slower than lower wavelengths. Thus, for $\Delta\lambda_0 \approx 2$ nm, the actual broadening of the pulse will be

$$\Delta\tau_m \approx 43 \text{ ps/km},$$

implying that because of material dispersion, the pulse will broaden by 43 ps after traversing through 1 km of silica fiber.

8.5.2 Waveguide dispersion

As discussed earlier, because of the dependence of refractive index on wavelength, different wavelengths take different amounts of time to propagate along the fiber, which is known as *material dispersion*. We will show here that even if n_1 and n_2 are independent of wavelength (i.e., there is no material dispersion), the group velocity of a particular mode will depend on the wavelength. This leads to what is known as the *waveguide dispersion*.

Recall that β can be approximated [see Eq. (8.32)] as

$$\beta = \frac{\omega}{c} n_2 \left[1 + \Delta\, b(V)\right].$$

We assume that n_1 and n_2 are independent of ω and calculate group velocity:

$$\frac{1}{v_g} = \frac{d\beta}{d\omega} = \frac{1}{c} n_2 \left[1 + \Delta\, b(V)\right] + \frac{\omega}{c} n_2 \Delta \frac{db}{dV} \frac{dV}{d\omega}. \tag{8.44}$$

Now,
$$V = \frac{2\pi}{\lambda_0} a \sqrt{n_1^2 - n_2^2} = \frac{\omega}{c} a \sqrt{n_1^2 - n_2^2}\ . \tag{8.45}$$

Thus,
$$\frac{dV}{d\omega} = \frac{V}{\omega}\ , \tag{8.46}$$

and

$$\begin{aligned} \frac{1}{v_g} &= \frac{n_2}{c}\left[1 + \Delta\, b(V)\right] + \frac{n_2}{c} \Delta\, V \frac{db}{dV} \\ &= \frac{n_2}{c}\left[1 + \Delta\, \frac{d}{dV}(bV)\right]. \end{aligned} \tag{8.47}$$

Thus, the time taken by a pulse to traverse length L of the fiber is given by

$$\tau = \frac{L}{v_g} = \frac{L}{c} n_2 \left[1 + \Delta\, \frac{d}{dV}(bV)\right]. \tag{8.48}$$

From Eq. (8.48), we see that even if n_1 and n_2 are independent of wavelength, the group velocity (and hence τ) will still depend on ω because of the dependence of b on V (see Fig. 8.3). This leads to what is known as the *waveguide dispersion*. Physically, waveguide dispersion arises due to the fact that the spot size depends

on the wavelength. For a source having a spectral width $\Delta\lambda_0$, the corresponding waveguide dispersion is given by

$$\Delta\tau_{\mathrm{w}} = \frac{d\tau}{d\lambda_0}\Delta\lambda_0 \cong \frac{L}{c} n_2 \Delta \frac{d^2}{dV^2}(bV)\frac{dV}{d\lambda_0}\Delta\lambda_0 . \tag{8.49}$$

From Eq. (8.45), we find that

$$\frac{dV}{d\lambda_0} = -\frac{V}{\lambda_0}. \tag{8.50}$$

Thus,
$$\Delta\tau_{\mathrm{w}} = -\frac{L n_2 \Delta}{c} f(V)\frac{\Delta\lambda_0}{\lambda_0}. \tag{8.51}$$

The corresponding dispersion coefficient is defined as

$$D_{\mathrm{w}} = \frac{\Delta\tau_{\mathrm{w}}}{L\Delta\lambda_0} = -\frac{n_2\Delta}{c\lambda_0} f(V), \tag{8.52}$$

where
$$f(V) \equiv V\frac{d^2}{dV^2}(bV). \tag{8.53}$$

For a step-index fiber, b as a function of V is a universal curve; therefore, the variation of $f(V)$ with V will also be universal. A convenient empirical formula for a step-index fiber is given by[9]

$$f(V) \approx 0.080 + 0.549(2.834 - V)^2 \quad 1.3 < V < 2.4. \tag{8.54}$$

Using the preceding equation, the empirical expression for waveguide dispersion in step-index fibers is

$$D_{\mathrm{w}} = -\frac{n_2\Delta}{3\lambda_0} \times 10^4 \left[0.080 + 0.549(2.834 - V)^2\right] \ \mathrm{ps/(nm \cdot km)}, \tag{8.55}$$

where λ_0 is measured in μm.

Thus, the total dispersion in the case of a single-mode optical fiber can be attributed to two types of dispersion—namely, material dispersion and waveguide dispersion. Indeed, it can be shown that in most single-mode fibers, the total dispersion coefficient D_{tot} is given to a good accuracy by the sum of material D_{m} and waveguide D_{w} dispersion coefficients. The material contribution

is given by Eq. (8.42), while the waveguide contribution for a step-index fiber is given by Eq. (8.55).

In the single-mode regime, the quantity within the bracket in Eq. (8.55) is usually positive; hence, the waveguide dispersion is negative. Since the sign of material dispersion depends on the operating wavelength region, it is possible that the two effects—namely, material dispersion and waveguide dispersion—cancel each other at a certain wavelength. Such a wavelength, which is a very important parameter of single-mode fibers, is referred to as the *zero-dispersion wavelength* λ_{ZD}. Single-mode step-index optical fibers that have a zero-dispersion wavelength around 1300 nm are known as *standard single-mode fibers* (SSMFs) or G-652 fibers. These fibers have been used extensively in optical communication systems.

In Fig. 8.5, we show the wavelength dependence of D_m, D_w, and D_{tot} for a typical SSMF with parameters as given in Example 8.2. The variation of D_m is calculated using Eqs. (8.42) and (8.43) and that of D_w using Eq. (8.55). The figure shows that the total dispersion passes through zero around $\lambda_0 = 1300$ nm.

8.5.3 Dispersion-shifted fibers

It may be noted from Eq. (8.55) that the waveguide dispersion depends on the fiber parameters, which can be selected to make the total dispersion zero at some desired wavelength. Because the lowest loss in an optical fiber occurs at a wavelength of 1550 nm and optical amplifiers are also available in the 1550-nm window, fiber designs can be modified to shift the zero-dispersion wavelength to the 1550-nm wavelength window. As an example, we consider a step-index single-mode fiber with parameters as

$$n_2 = 1.444,\ \Delta = 0.0075, \text{ and } a = 2.3\ \mu\text{m, giving } V = 2.556/\lambda_0,$$

where λ_0 is measured in μm. Using the preceding parameters in Eq. (8.55), at $\lambda_0 =$ 1550 nm one obtains

$$D_w \approx -20 \text{ ps/(nm·km)}.$$

On the other hand, using Eqs. (8.42) and (8.43), the material dispersion at this wavelength is

$$D_m \approx +20 \text{ ps/(nm·km)}.$$

This shows that for the considered fiber parameters, the total dispersion is zero at $\lambda_0 = 1550$ nm. The corresponding variation of D_m, D_w, and D_{tot} with wavelength is shown in Fig. 8.6. The preceding example shows that one can shift the zero dispersion wavelength to a suitable value by changing the fiber parameters. Such fibers are referred to as *dispersion-shifted fibers* (DSFs) when dispersion is zero around 1550 nm and *nonzero-dispersion-shifted fibers* (NZDSFs) when dispersion is finite but small around 1550 nm.

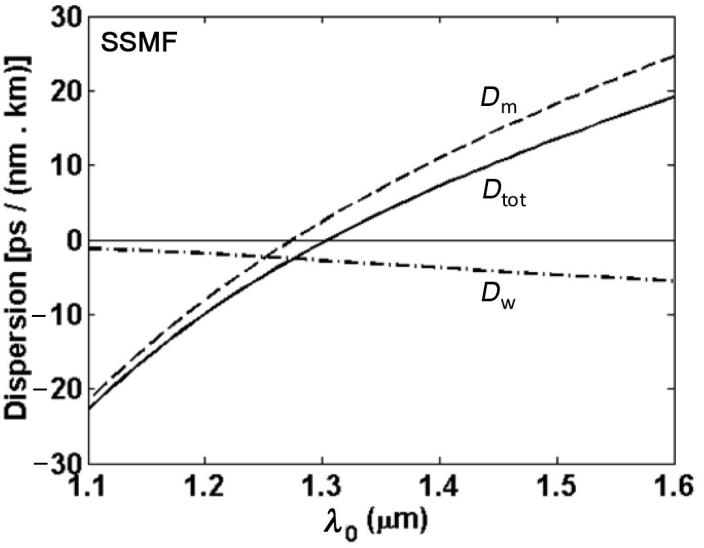


Figure 8.5 The wavelength dependence of D_m, D_w, and D_{tot} for a typical conventional SSMF with parameters given in Example 8.2, showing $\lambda_{ZD} \approx 1300$ nm. (Calculations courtesy of Ms. Triranjita Srivastava.)

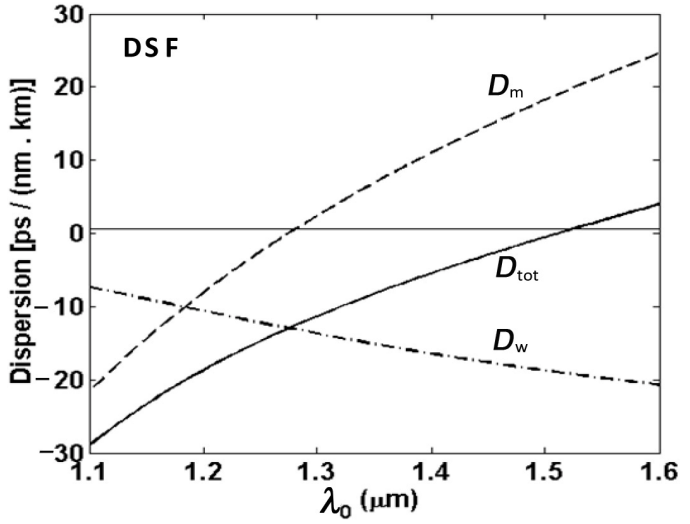


Figure 8.6 The wavelength dependence of D_m, D_w, and D_{tot} for a typical dispersion-shifted fiber with parameters as $n_2 = 1.444$, $\Delta = 0.0075$, and $a = 2.3$ μm. The zero-dispersion wavelength is around 1550 nm. (Calculations courtesy of Ms. Triranjita Srivastava.)

8.6 Polarization Characteristics of Single-Mode Fibers

The discussion and the numerical results presented in the preceding sections use LP_{lm} modes, which are extremely accurate for practical optical fibers, as $n_1 \approx n_2$.

In this approximation, various modes are linearly polarized. The exact modes, obtained by satisfying the exact boundary conditions at the core–cladding interface, however, are not perfectly linearly polarized. In the following section, we briefly discuss the exact vector modes of a step-index fiber.

8.6.1 Exact vector modes of a step-index fiber

In a step-index fiber, both core and cladding regions are homogeneous, and hence, as discussed in Section 2.2, the field components E_z and H_z in each region will satisfy the following wave equations:

$$\nabla^2 E_z = \varepsilon \mu \frac{\partial^2 E_z}{\partial t^2}, \tag{8.56a}$$

and

$$\nabla^2 H_z = \varepsilon \mu \frac{\partial^2 H_z}{\partial t^2}. \tag{8.56b}$$

Keeping the cylindrical geometry of the fiber in mind, we solve these equations in cylindrical coordinates (r, ϕ, and z); z being parallel to the fiber axis and also taken as the direction of propagation. We first solve the preceding equations for E_z and H_z. Since the refractive index of the fiber does not depend on the z coordinate, we can assume E_z and H_z to be of the form

$$\begin{aligned} \mathbf{E}(r,\phi,z,t) &= \mathbf{E}(r,\phi)\,\mathrm{e}^{i(\omega t-\beta z)} \\ \mathbf{H}(r,\phi,z,t) &= \mathbf{H}(r,\phi)\,\mathrm{e}^{i(\omega t-\beta z)}, \end{aligned} \tag{8.57}$$

where β represents the propagation constant of the mode. Substituting the preceding expressions into Maxwell's Eqs. (2.11) and (2.12) written in cylindrical coordinates, we can show[4,10] that the other field components E_r, E_ϕ, H_r, and H_ϕ are given by

$$\begin{aligned} E_r &= -\frac{i}{\alpha^2}\left(\beta\frac{\partial E_z}{\partial r} + \omega\mu\frac{1}{r}\frac{\partial H_z}{\partial \phi}\right), \\ E_\phi &= -\frac{i}{\alpha^2}\left(\beta\frac{1}{r}\frac{\partial E_z}{\partial \phi} - \omega\mu\frac{\partial H_z}{\partial r}\right), \\ H_r &= -\frac{i}{\alpha^2}\left(\beta\frac{\partial H_z}{\partial r} - \omega\varepsilon\frac{1}{r}\frac{\partial E_z}{\partial \phi}\right), \\ H_\phi &= -\frac{i}{\alpha^2}\left(\beta\frac{1}{r}\frac{\partial H_z}{\partial \phi} + \omega\varepsilon\frac{\partial E_z}{\partial r}\right), \end{aligned} \tag{8.58}$$

where $\alpha^2 = k_0^2 n^2(r) - \beta^2$ and $k_0^2 = \omega^2 \varepsilon_0 \mu_0$. Thus, all of the field components are known once E_z and H_z are known. Substituting E_z and H_z from Eq. (8.57) into Eqs. (8.56a) and (8.56b), one can easily show that $E_z(r,\phi)$ and $H_z(r,\phi)$ satisfy the following equation:

$$\frac{\partial^2 \psi}{\partial r^2} + \frac{1}{r}\frac{\partial \psi}{\partial r} + \frac{1}{r^2}\frac{\partial^2 \psi}{\partial \phi^2} + \left[k_0^2 n^2(r) - \beta^2\right]\psi = 0, \tag{8.59}$$

where $\psi(r,\phi)$ represents $E_z(r,\phi)$ or $H_z(r,\phi)$. As discussed in Section 8.3, Eq. (8.59) can be solved by using the method of separation of variables, and for the guided modes, the well-behaved solutions are given by

$$\psi(r,\phi) = \begin{cases} \dfrac{A}{J_l(U)} J_l\left(\dfrac{Ur}{a}\right)\begin{bmatrix}\cos l\phi \\ \sin l\phi\end{bmatrix} & r < a, \\ \dfrac{A}{K_l(W)} K_l\left(\dfrac{Wr}{a}\right)\begin{bmatrix}\cos l\phi \\ \sin l\phi\end{bmatrix} & r > a, \end{cases} \tag{8.60}$$

where A is a constant, and U and W are the same, as defined in Eq. (8.17). In Eq. (8.60), we have already implemented the continuity of E_z and H_z at $r = a$.

Having determined the E_z and H_z components, the other field components can be determined by using Eq. (8.58). It should be mentioned here that the ϕ dependence of E_r, E_ϕ, H_r, and H_ϕ should also be of the $\cos l\phi$ or $\sin l\phi$ type. This requires that if the ϕ dependence of E_z is the $\cos l\phi$ ($\sin l\phi$) type, the ϕ dependence of H_z should be the $\sin l\phi$ ($\cos l\phi$) type [see Eq. (8.58)]. Thus,

$$E_z(r,\phi) = \begin{cases} \dfrac{A}{J_l(U)} J_l\left(\dfrac{Ur}{a}\right)\begin{bmatrix}\cos l\phi \\ \sin l\phi\end{bmatrix} & r < a, \\ \dfrac{A}{K_l(W)} K_l\left(\dfrac{Wr}{a}\right)\begin{bmatrix}\cos l\phi \\ \sin l\phi\end{bmatrix} & r > a, \end{cases} \tag{8.61}$$

$$H_z(r,\phi) = \begin{cases} \dfrac{B}{J_l(U)} J_l\left(\dfrac{Ur}{a}\right)\begin{bmatrix}\sin l\phi \\ \cos l\phi\end{bmatrix} & r < a, \\ \dfrac{B}{K_l(W)} K_l\left(\dfrac{Wr}{a}\right)\begin{bmatrix}\sin l\phi \\ \cos l\phi\end{bmatrix} & r > a. \end{cases} \tag{8.62}$$

Substituting $E_z(r,\phi)$ and $H_z(r,\phi)$ in Eq. (8.58), one obtains E_r, E_ϕ, H_r, and H_ϕ in the core as well as the cladding regions. If we make E_ϕ and H_ϕ continuous at $r = a$, we obtain the following two equations:[4]

$$A\beta\left[\frac{1}{U^2}+\frac{1}{W^2}\right]l = -B\omega\mu_0\left[\frac{J_l'(U)}{UJ_l(U)}+\frac{K_l'(W)}{WK_l(W)}\right], \tag{8.63}$$

and

$$A\omega\varepsilon_0\left[n_1^2\frac{J_l'(U)}{UJ_l(U)}+n_2^2\frac{K_l'(W)}{WK_l(W)}\right] = -B\beta\left[\frac{1}{U^2}+\frac{1}{W^2}\right]l. \tag{8.64}$$

Dividing Eq. (8.64) by Eq. (8.63), we obtain

$$\left[\frac{J_l'(U)}{UJ_l(U)}+\frac{K_l'(W)}{WK_l(W)}\right]\left[n_1^2\frac{J_l'(U)}{UJ_l(U)}+n_2^2\frac{K_l'(W)}{WK_l(W)}\right] = \frac{\beta^2}{k_0^2}\left[\frac{1}{U^2}+\frac{1}{W^2}\right]^2 l^2. \tag{8.65}$$

The preceding transcendental equation is the exact eigenvalue equation whose solutions give the propagation constants of various guided modes of a step-index fiber. The modes are, in general, hybrid modes consisting of all of the six field components except for $l = 0$. In the case where $l = 0$, Eq. (8.65) is further divided into the following two equations:

TE modes:
$$\frac{J_0'(U)}{UJ_0(U)} + \frac{K_0'(W)}{WK_0(W)} = 0, \tag{8.66}$$

TM modes:
$$n_1^2\frac{J_0'(U)}{UJ_0(U)} + n_2^2\frac{K_0'(W)}{WK_0(W)} = 0. \tag{8.67}$$

Note that when Eq. (8.66) is satisfied, from Eq. (8.63), $A = 0$, implying that $E_z = 0$. This means that for such modes, the longitudinal component of the electric field is zero, and hence, the modes are known as *transverse electric* (*TE*) *modes*. Similarly, when Eq. (8.67) is satisfied, from Eq. (8.64), $B = 0$, implying that $H_z = 0$. Thus, for such modes, the longitudinal component of the magnetic field is zero, and hence, the modes are known as *transverse magnetic* (*TM*) *modes*.

For $l \geq 1$, as mentioned earlier, the modes are hybrid modes consisting of all of the six field components. The propagation constants of these modes are obtained by solving Eq. (8.65), which results in two types of modes designated as HE and EH modes. In the following section, we derive the eigenvalue equations of HE and EH modes in the weakly guiding approximation ($\Delta << 1$).

8.6.1.1 Eigenvalue equations of HE and EH modes for $\Delta << 1$

We first put β^2 in terms of U^2 and W^2 in Eq. (8.65). Using the definitions of U and W [Eq. (8.17)], we can write

$$\frac{k_0^2 n_1^2 - \beta^2}{U^2} = \frac{\beta^2 - k_0^2 n_2^2}{W^2},$$

i.e,

$$\frac{\beta^2}{k_0^2}\left[\frac{1}{U^2} + \frac{1}{W^2}\right] = \left[\frac{n_1^2}{U^2} + \frac{n_2^2}{W^2}\right]. \tag{8.68}$$

Using the preceding expression in Eq. (8.65), the eigenvalue equation can be recast as

$$\left[\frac{J_l'(U)}{UJ_l(U)} + \frac{K_l'(W)}{WK_l(W)}\right]\left[n_1^2 \frac{J_l'(U)}{UJ_l(U)} + n_2^2 \frac{K_l'(W)}{WK_l(W)}\right] = l^2 \left[\frac{1}{U^2} + \frac{1}{W^2}\right]\left[\frac{n_1^2}{U^2} + \frac{n_2^2}{W^2}\right]. \tag{8.69}$$

If we take the limit $n_1 \to n_2$, the preceding equation gives the following pair of equations:

$$\left[\frac{J_l'(U)}{UJ_l(U)} + \frac{K_l'(W)}{WK_l(W)}\right] = \pm l \left[\frac{1}{U^2} + \frac{1}{W^2}\right]. \tag{8.70}$$

The equation resulting in the positive term on the right-hand side in the preceding equation belongs to the EH mode, and the equation resulting in the negative term on the right-hand side belongs to the HE mode. Using the following recurrence relations of Bessel functions given in Eqs. (8.20a) and (8.20b), Eq. (8.70) can be rewritten as

EH modes:
$$\frac{J_{l+1}(U)}{UJ_l(U)} = -\frac{K_{l+1}(W)}{WJ_l(W)}, \tag{8.71}$$

HE modes:
$$\frac{J_{l-1}(U)}{UJ_l(U)} = +\frac{K_{l-1}(W)}{WK_l(W)}. \tag{8.72}$$

It may be mentioned here that for a given value of l, there may be several possible solutions, and the m'th solution of each equation is denoted as the EH_{lm} and HE_{lm} mode, respectively.

8.6.1.2 Cutoff conditions of the various modes

As discussed earlier, the cutoff V values of various guided modes can be obtained from the eigenvalue Eq. (8.65) by substituting the values of Bessel functions in the limit: $\beta = k_0 n_2$, i.e., $W = 0$ and $U = V_c$. The cutoffs of different modes are the various roots[10] of

(i) $$J_0(V_c) = 0, \text{ TE}_{0m} \text{ and TM}_{0m} \text{ modes } (l = 0), \tag{8.73}$$

(ii) $$J_1(V_c) = 0, \text{ HE}_{1m} \text{ mode } (l = 1), \tag{8.74}$$

$$\left(\frac{n_1^2}{n_2^2} + 1\right) J_{l-1}(V_c) = \frac{V_c}{l-1} J_l(V_c), \text{ HE}_{lm} \text{ mode } (l \geq 2). \tag{8.75a}$$

In the limit $n_1 \to n_2$, the preceding equation reduces to

$$J_{l-2}(V_c) = 0, \tag{8.75b}$$

(iii) $$J_l(V_c) = 0, \; V_c \neq 0 \text{ EH}_{lm} \text{ mode.} \tag{8.76}$$

Using the preceding equations, we arrange below the first few modes according to the increasing cutoff V values:

0.000	HE_{11}
2.4048	TE_{01}, TM_{01}, and HE_{21}
3.8317	HE_{12} and EH_{11}.

(Note that the cutoff of the HE_{21} mode is obtained in the limit $n_1 \to n_2$, i.e., from Eq. (8.75b). The exact value depends on n_1/n_2 [see Eq. (8.75a)] and is slightly higher than 2.4048.) In the limit $n_1 \to n_2$, these vector modes are converted to LP_{lm} modes; the correspondence between the two modes is given by[4]

$$\begin{aligned} &LP_{0m} = HE_{1m} \\ &LP_{1m} = TE_{0m}, TM_{0m}, HE_{2m} \\ &LP_{lm} = HE_{l+1,m}, EH_{l-1,m} \; (l \geq 2). \end{aligned} \tag{8.77}$$

The most important observation here is that the HE_{11} mode has zero cutoff and hence exists at all frequencies. As the V parameter increases and approaches 2.4048, the next higher-order modes (TE_{01}, TM_{01}) will also be supported. Thus, the single-mode region is again given by the condition $0 < V < 2.4048$. If this condition is satisfied, the fiber will support only one mode—namely, the HE_{11} mode.

The so-called single-mode fibers, in fact, support two orthogonally polarized modes, propagating with propagation constants β_x and β_y, respectively. Figure 8.7 schematically depicts the polarization states of these two fundamental vector modes—namely, HE_{11}^{x} and HE_{11}^{y}.

It may be noted that that the HE_{11}^{x} or HE_{11}^{y} mode is predominantly *x* or *y* polarized, having also a tiny *y* or *x* component. Under ideal conditions, circular-core optical fibers have no birefringence, i.e., ($\beta_x = \beta_y$). However, practical fibers do not have a perfectly circular core. As a result, the two modes have slightly different propagation constants ($\beta_x \neq \beta_y$), and the fiber becomes birefringent. The birefringence is measured in terms of *B* or Δn_{eff}, defined as

$$B = \beta_x - \beta_y \qquad \text{or} \qquad \Delta n_{\mathrm{eff}} = B / k_0 .$$

Apart from core ellipticity, many other factors can introduce birefringence in the fiber. In Section 8.6.2, we discuss the various origins of birefringence in optical fibers.

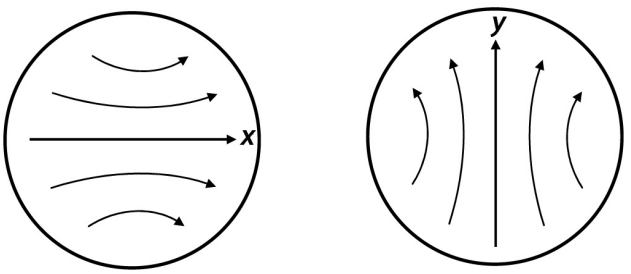
8.6.2 Origins of birefringence in optical fibers

The various reasons that an optical fiber may become birefringent are shown in Fig. 8.8 and are discussed in the remainer of this chapter.

8.6.2.1 Core ellipticity

The geometrical anisotropy due to core ellipticity [Fig. 8.8(a)] introduces a linear birefringence in the fiber. For fibers with small core ellipticity, the birefringence near the first higher-mode cutoff is approximately given by[11]

$$B_{\mathrm{ce}} = (\beta_x - \beta_y) \approx 0.2 k_0 \left(\frac{a}{b} - 1 \right) (\Delta n)^2 , \tag{8.78}$$



HE_{11}^{x} HE_{11}^{y}

Figure 8.7 Schematic of polarization states for orthogonally polarized HE_{11}^{x} and HE_{11}^{y} modes.

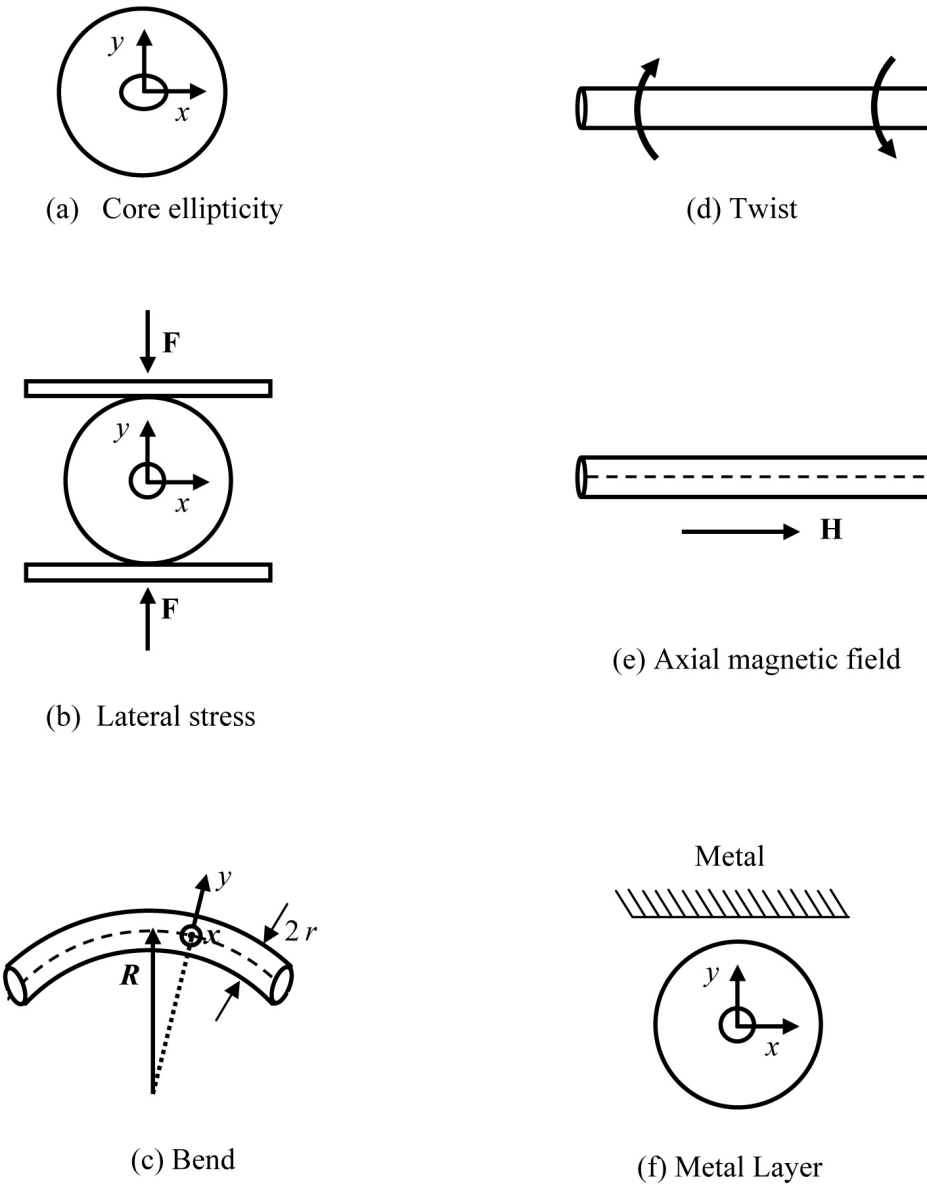


Figure 8.8 Various mechanisms that produce birefringence in an optical fiber.

where $k_0 = 2\pi / \lambda_0$ (λ_0 being the free-space wavelength), Δn is the core and cladding refractive index difference, and a and b represent the semi-major and semi-minor axes of the core ellipse. In Eq. (8.78), β_x and β_y represent the propagation constants of the x- and y-polarized modes, where the x and y axes are taken parallel to the major and minor axis, respectively.

Equation (8.78) indicates that the birefringence is higher in fibers with larger core ellipticity and higher core and cladding refractive index difference. A positive value of B_{ce} means that the *x* component travels as a slow component. The corresponding difference between the effective indices of the slow and fast components is thus given by

$$\Delta n_{\text{eff}} = \frac{B_{\text{ce}}}{k_0} \approx 0.2\left(\frac{a}{b}-1\right)(\Delta n)^2 . \tag{8.79}$$

Example 8.5 We consider a single-mode optical fiber with $\Delta n = 0.005$ and core ellipticity of 1%, i.e., $(a/b-1) = 0.01$. Equation (8.79) gives $\Delta n_{\text{eff}} = 5\times10^{-8}$. Core ellipticity is also responsible for asymmetrical stresses developed in the fiber core. These stresses generate an extra birefringence, which is discussed in the next chapter.

8.6.2.2 Lateral stress

An asymmetrical lateral stress on the fiber also produces linear birefringence in the fiber via a strain-optic effect. If the fiber is pressed (for instance, along the *y* axis) between two parallel plates [see Fig. 8.8(b)], the generated birefringence is given by[11,12]

$$B_{\text{s}} = (\beta_x - \beta_y) = -k_0 h \frac{F}{r}, \tag{8.80}$$

with

$$h = \frac{2 n^3}{\pi}\frac{(1+\nu)}{E}\left(p_{11}-p_{12}\right). \tag{8.81}$$

Here, n is the average refractive index of the fiber core and cladding, p_{11} and p_{12} are the components of the strain optical tensor, E is the Young's modulus, and ν is the Poisson's ratio of the fiber material. Further, r represents the outer radius (including the cladding) of the fiber, which for a standard fiber is 62.5 μm, and F represents the lateral force per unit length of the fiber in kg/m.

Taking n = 1.46, p_{11} = 0.12, p_{12} = 0.27, $E = 7.75 \times 10^9$ kg/m^2, and ν = 0.17 for fused silica, the difference between the effective indices of the lateral-stress-induced slow and fast components is

$$\Delta n_{\text{eff}} = \frac{B_{\text{s}}}{k_0} = 4.48\times10^{-11}\left(\frac{F}{r}\right). \tag{8.82}$$

Example 8.6 Substituting F = 1 gm/cm, i.e., 0.1 kg/m, and $r = 62.5\times10^{-6}$ m in Eq. (8.82), one obtains $\Delta n_{\text{eff}} \approx 7\times10^{-8}$.

8.6.2.3 Bending

Bending of a fiber introduces different types of stresses in the outer and inner portion of the fiber, leading to a linear birefringence, which for silica fibers is given by[13]

$$B_{\mathrm{b}} = (\beta_x - \beta_y) = -k_0 \frac{n^3}{4}(1+\nu)\left(p_{11} - p_{12}\right)\left(\frac{r}{R}\right)^2 \text{ rad/m}, \qquad (8.83)$$

where R represents the bend radius. In the preceding equation, the x axis is taken perpendicular to the plane of bending and the y axis in the plane of bending [Fig. 8.8(c)]. Substituting $n = 1.46$, $p_{11} = 0.12$, $p_{12} = 0.27$, and $\nu = 0.17$ in Eq. (8.83), for fused silica, the difference between the effective indices of the bending-induced slow and fast components is

$$\Delta n_{\mathrm{eff}} = \frac{B_{\mathrm{b}}}{k_0} = 0.136\left(\frac{r}{R}\right)^2. \qquad (8.84)$$

Example 8.7 Substituting R = 10 cm, i.e., 0.1 m, and $r = 62.5\times10^{-6}$ m in Eq. (8.84), one obtains $\Delta n_{\mathrm{eff}} \approx 5.3\times10^{-8}$. Bending-induced birefringence is used to fabricate in-line fiber optic wave plates to control the polarization state of the guided light, as discussed in the next chapter.

8.6.2.4 Twists

A twist [Fig. 8.8(d)] in the fiber produces a circular birefringence, given by[14]

$$B_{\mathrm{t}} = (\beta_{\mathrm{rc}} - \beta_{\mathrm{lc}}) = g\tau \text{ rad/m}, \qquad (8.85)$$

where β_{rc} and β_{lc} represent the propagation constants of the right-circularly and left-circularly polarized modes, respectively, $g = -n^2/2\left(p_{11} - p_{12}\right) \approx 0.16$ for silica fibers, and τ is the twist rate (in rad/m). Thus, a linearly polarized light will be rotated along a twisted fiber with a rotation rate given by

$$\alpha = \frac{1}{2} g\tau \text{ rad/m}. \qquad (8.86)$$

A positive value of g means that if the twist is counterclockwise (with respect to an observer looking against the direction of propagation), the previously mentioned rotation will also be counterclockwise (i.e., the left-circular component will be faster). Thus, the difference between the effective indices of the twist-induced slow and fast components is given by

$$\Delta n_{\text{eff}} = (n_{\text{rc}} - n_{\text{lc}}) = \frac{B_{\text{b}}}{k_0} = \frac{\lambda_0}{2\pi} g\tau \,. \tag{8.87}$$

Example 8.8 A twist of 1 turn/m in a fiber at $\lambda_0 = 1.5\,\mu\text{m}$ will produce

$$\Delta n_{\text{eff}} = \frac{1.5\times10^{-6}}{2\pi} \times 0.16 \times 2\pi = 2.4\times10^{-7} \,.$$

8.6.2.5 Magnetic field

A longitudinal (i.e., axial) magnetic field **H** [Fig. 8.8(e)] produces a circular birefringence in the fiber via the Faraday effect and is given by[15]

$$B_{\text{H}} = \beta_{\text{rc}} - \beta_{\text{lc}} = \frac{2V_{\text{d}}}{l} \int \mathbf{H} \bullet \mathbf{dl} \,, \tag{8.88}$$

where l represents the length of the fiber, and the line integral is taken along the direction of propagation. Further, V_{d} is the Verdet constant of the material and is equal to 4.7×10^{-6} rad/A for silica fibers.

It should be noted that unlike in the previous cases, the sign of birefringence produced by an axial magnetic field depends on the direction of propagation relative to **H**. If the wave is propagating along **H**, then $\beta_{\text{rc}} > \beta_{\text{lc}}$ (left handed), and if the wave is propagating opposite to **H**, then $\beta_{\text{rc}} < \beta_{\text{lc}}$ (right handed). An important property of the Faraday effect is that it is nonreciprocal, which means that if light propagating in one direction undergoes rotation θ, light propagating in the opposite direction undergoes rotation $-\theta$. This property is exploited in realizing important polarization components such as optical isolators and circulators. In a constant magnetic field of magnitude H, birefringence produced by an axial magnetic field is given by

$$\left|B_{\text{H}}\right| = 2V_{\text{d}} H \text{ rad/m.} \tag{8.89}$$

Thus, an axial magnetic field of magnitude H A/m will rotate the plane of polarization of the guided wave with a rate of

$$\alpha = V_{\text{d}} H \text{ rad/m.} \tag{8.90}$$

Example 8.9 An axial magnetic field of magnitude 1 A/m will generate a birefringence of 9.4×10^{-6} rad/m; thus, the polarization ellipse of a propagating polarized beam will rotate with a rate of 4.7×10^{-6} rad/m along the fiber. One of the important applications of the Faraday effect is in the measurement of electric currents using optical fibers, as discussed in the next chapter.

8.6.2.6 Metal layer near the fiber core

A dielectric–metal interface supports a TM-polarized surface wave (i.e., an electric field perpendicular to the interface), known as a surface plasmon polariton (SPP). Thus, if an optical fiber is side polished up to near its core and a metal layer is deposited on it [Fig. 8.8(f)], the fiber becomes birefringent. The TE (x-polarized) and the TM (y-polarized) modes propagate with different propagation constants and loss coefficients. Depending on the metal, the distance (from the core), and the thickness of the metal layer, the TM mode may be highly lossy due to coupling between the fiber mode and the SPP supported by the dielectric–metal interface. This mechanism is used to realize efficient in-line fiber polarizers and many biochemical sensors.

It may be mentioned that a deliberate and controlled birefringence introduced in a fiber can be used to realize several in-line fiber components and devices. In the next chapter, we will discuss some applications of birefringence in optical fibers. On the other hand, a random and uncontrolled birefringence in a fiber causes many problems, as is discussed in Chapter 10.

References

1. T. Miya, Y. Terunama, T. Hosaka, and T. Miyasita, "An ultimate low loss single mode fiber at 1.55 μm," *Electron. Lett.* **15**, 106–108 (1979).
2. A. Ghatak and K. Thyagarajan, *Optical Electronics*, Cambridge University Press, New York (1989).
3. A. Ghatak and K. Thyagarajan, *Introduction to Fiber Optics*, Cambridge University Press, Cambridge, UK (1998).
4. K. Okamota, *Fundamentals of Optical Waveguides*, Academic Press, San Diego, CA (2000).
5. H. D. Rudolph and E. G. Neumann, "Approximation for the eigenvalues of the fundamental mode of a step index glass fiber waveguide," *Nachrichtentech*. Z. **4**, 328–329 (1976).
6. D. Marcuse, "Loss analysis of single mode fiber splices," *Bell Syst. Tech. J*. **56**, 703–718 (1977).
7. A. Ghatak, *Optics*, McGraw-Hill, New York (2009).
8. U. C. Paek, G. E. Peterson, and A. Carnevale, "Dispersionless single mode light guides with α index profiles," *Bell Syst. Tech. J.* **60**, 583–598 (1981).
9. D. Marcuse, "Interdependence of waveguide and material dispersion," *Appl. Opt*. **18**, 2930–2932 (1979).
10. D. Marcuse, *Light Transmission Optics*, 2nd ed., Van Nostrand Reinhold Co., New York (1982).
11. S. C. Rashleigh, "Origins and control of polarization effects in single-mode fibers," *J. Lightw. Technol.* **1**, 312–331 (1983).

12. D. N. Payne, A. J. Barlow, and J. J. R. Hansen, "Development of low and high-birefringence optical fibers," *IEEE J. Quantum Electron.* **18**, 477–488 (1982).
13. R. Ulrich, S. C. Rashleigh, and W. Eickhoff, "Bending-induced birefringence in single-mode fibers," *Opt. Lett.* **5,** 273–275 (1980).
14. R. Ulrich and A. Simon, "Polarization optics of twisted single-mode fibers," *Appl. Opt.* **18**, 2241–2251 (1979).
15. A. M. Smith, "Polarization and magneto-optical properties of single-mode fibers," *Appl. Opt.* **17**, 52–56 (1978).

Bibliography

Ghatak, A., I. C. Goyal, and R. Varshney, *FIBER OPTICA: A Software for Characterizing Fiber and Integrated-Optic Waveguides*, Viva Books, New Delhi (1999).

Gloge, D., "Weakly guiding fibers," *Appl. Opt.* **10**, 2252–2258 (1971).

Marcuse, D., "Gaussian approximation of the fundamental modes of graded index fibers," *J. Opt. Soc. Am.* **68**, 103–109 (1978).

Mynbaev, D. K. and L. L. Scheiner, *Fiber-Optic Communications Technology*, Prentice Hall, Upper Saddle River, NJ (2001).

Snyder, A. W. and J. D. Love, *Optical Waveguide Theory*, Chapman & Hall, London (1983).

Chapter 9
Birefringence in Optical Fibers: Applications

9.1 Introduction

As discussed in the previous chapter, the so-called single-mode fibers in fact support two modes simultaneously, which are orthogonally polarized. In an ideal circular-core fiber, these two modes will propagate with the same phase velocity; however, practical fibers are not perfectly circularly symmetric. As a result, the two modes propagate with slightly different phase and group velocities. Furthermore, environmental factors such as bend, twist, and anisotropic stress also produce birefringence in the fiber, the direction and magnitude of which keep changing with time due to changes in the ambient conditions such as temperature. These factors also couple energy from one mode to the other mode of the fiber, creating problems in practical applications.

The first issue we discuss is that as the magnitude of the birefringence mentioned earlier keeps changing randomly with time due to fluctuations in the ambient conditions, the output SOP also keeps fluctuating with time. The change in the output SOP is of little consequence in applications where the detected light is not sensitive to the polarization state. However, in many applications such as fiber optic interferometric sensors, coupling between optical fibers and integrated optic devices, coherent communication systems, and so on, the output SOP must remain stable. Secondly, since each mode propagates with a different group velocity, the so-called polarization mode dispersion (PMD) can limit the ultimate bandwidth of a single-mode optical communication system.

We have already discussed the detrimental aspects of the random birefringence present in practical fibers. However, a controlled and deliberate birefringence introduced in the fiber may be used to realize several in-line fiber optical components and devices. In this chapter, we will discuss this aspect, i.e., the applications of the controlled and deliberate birefringence introduced in the fiber. The detrimental aspect—namely, the PMD—will be discussed in the next chapter.

One of the most important applications of the deliberately introduced birefringence in optical fibers is the development of high-birefringence (Hi-Bi) fibers, which are realized by introducing a permanent high birefringence into the

fiber during the fabrication stage itself. Such fibers can maintain the SOP of the incident light over large distances and hence are also known as polarization-maintaining fibers (PMFs). In the next section, we will discuss different types of PMFs, their polarization characteristics, and their applications. Devices using controlled birefringence in conventional single-mode fibers are then discussed.

9.2 Polarization-Maintaining Fibers

Polarization-maintaining fibers can be broadly divided into the following two categories: (i) high-birefringence (Hi-Bi) fibers and (ii) single-polarization single-mode (SPSM) fibers.

9.2.1 High-birefringence (Hi-Bi) fibers

In the case of highly birefringent fibers, the propagation constants of the two orthogonally polarized modes are made quite different from each other so that the coupling between the two modes is greatly reduced. Thus, if the light is coupled to only one of the polarized modes, most of the light remains in the same polarized mode; hence, the SOP of the propagating light is maintained along the fiber. The polarization-holding capacity of a birefringent fiber is measured in terms of its beat length L_b, which is defined by

$$L_{\mathrm{b}} = \frac{2\pi}{\beta_x - \beta_y} = \frac{\lambda_0}{\Delta n_{\mathrm{eff}}}, \tag{9.1}$$

where β_x and β_y represent the propagation constants of the two orthogonally polarized (for instance, along the x and y directions) modes, and Δn_{eff} represents the difference between the effective indices seen by the x- and y-polarized modes at wavelength λ_0. A lower value of L_{b} means a higher value of $\beta_x - \beta_y$ and hence corresponds to a fiber with high polarization-holding capacity. Physically, L_{b} represents the distance along the fiber in which the phase difference between the two fundamental modes becomes 2π. Thus, if light is coupled to both fundamental modes of such a fiber, the SOP will be repeated after each distance L_{b} along the fiber, as shown in Fig. 9.1. The typical values of L_{b} for Hi-Bi fibers lie between 1 and 2 mm.

Cross talk (CT) is another important parameter of a Hi-Bi fiber that indicates its practical polarization-holding capacity. CT is a measure of the power coupled to the orthogonal polarization due to random coupling along the fiber when power is launched into one of the polarized modes at the input end. For example, if P_x represents the power launched into the x-polarized mode at the input end and P_y represents the power detected in the y-polarized mode at the output end, then the CT is defined as

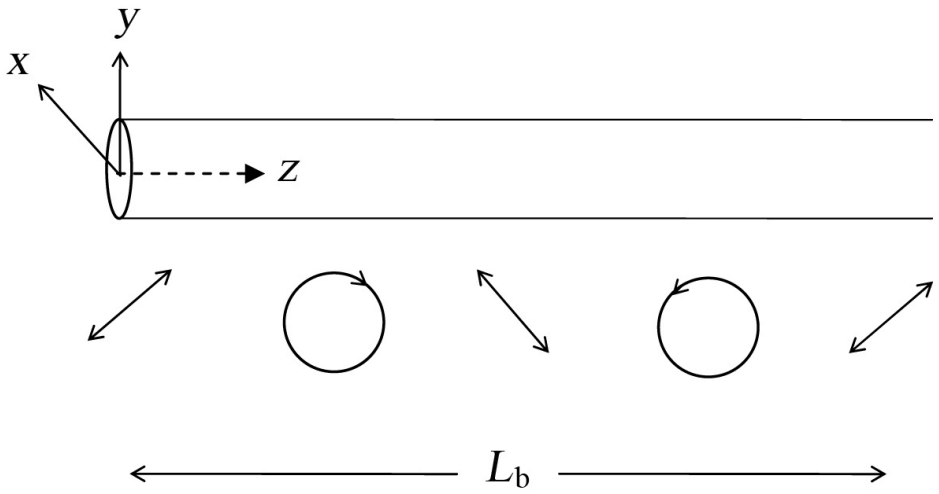


Figure 9.1 Evolution of the polarization state of light guided along a birefringent fiber when the *x*- and *y*-polarized modes are equally excited.

$$\mathrm{CT} = 10\log\left(\frac{P_y}{P_x}\right). \tag{9.2}$$

The unit of CT is decibels (dB), and it has a negative value. A higher absolute value of CT represents better polarization-holding capacity of the fiber.

Coupling between the two orthogonally polarized modes is also sometimes measured in terms of what is known as the *mode-coupling parameter* κ of a Hi-Bi fiber, which is defined as

$$\frac{P_y}{P_x} = \tanh(\kappa L), \tag{9.3}$$

L being the length of the fiber. The various types of popular Hi-Bi fibers are discussed in the remainder of Section 9.2.1.

9.2.1.1 Elliptical-core fibers

Elliptical-core fibers have an elliptical core embedded in a circular cladding, as shown in Fig. 9.2. The elliptical core of such fibers creates both geometrical anisotropy and asymmetrical stress in the core. As a result, the propagation constants β_x and β_y of the two fundamental modes polarized along the major axis (x direction) and minor axis (y direction), respectively, become different. The total birefringence is thus the sum of the geometrical birefringence B_g (due to the noncircular shape of the core) and stress-induced birefringence B_s (due to the asymmetrical stress produced by the elliptical core).

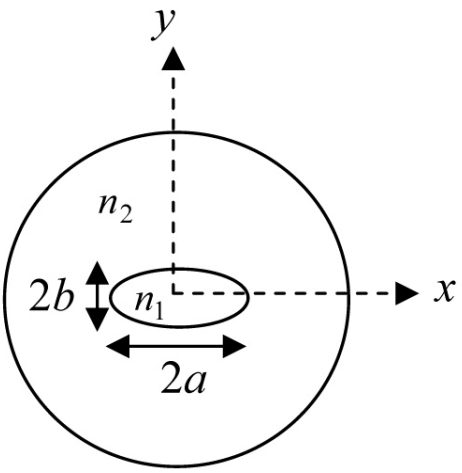


Figure 9.2 Transverse cross section of an elliptical core fiber.

9.2.1.1.1 Geometrical birefringence B_g

The geometrical birefringence of such fibers depends on the aspect ratio a/b of the core as well as on the refractive index difference $\Delta n = (n_1 - n_2)$ between the core and the cladding. In the vicinity of the first higher-mode cutoff, the birefringence for fibers with small core ellipticities ($a/b < 1.2$) is approximately given by[1]

$$B_g = (\beta_x - \beta_y) = 0.2k_0\left(\frac{a}{b} - 1\right)(\Delta n)^2 . \tag{9.4}$$

Equation (9.4) shows that the birefringence of such fibers can be increased by increasing either the core ellipticity or Δn. However, it should be mentioned that neither a/b nor Δncan be increased beyond a certain limit due to the following:

(i) The birefringence is not a sensitive function of a/b for $a/b \gg 2$; it saturates to a value that is approximately given by (as obtained by curve fitting from Fig. 1 of Ref. 2)

$$B_g \simeq 0.32(\Delta n)^2 . \tag{9.5}$$

(ii) For very large values of Δn, the core dimensions required for single-mode operation of the fiber become extremely small, and the fabrication of the fiber becomes very difficult. This is why such fibers were initially not considered of much use. However, the shift to longer wavelengths for long-distance communication systems has generated a great deal of interest in such fibers.

The exact calculation of B_g in elliptical-core fibers is quite difficult, as the wave equation must be solved in elliptical coordinates, and the eigenvalue equation so obtained is in the form of an infinite determinant.[3] Dyott et al.[4] obtained the birefringence of different elliptical-core fibers by truncating the infinite determinant to a finite order. Numerical techniques such as the finite-element method[5] and the point-matching method[6] have also been used by various authors; however, these approaches involve time-consuming numerical calculations. A number of approximate methods[7–9] have also been reported, among which the one proposed by Kumar et al.[9] is relatively simple and gives reasonably accurate values of B_g in the region of practical interest. This approach, which approximates the elliptical core by an equivalent rectangular core, is discussed in the following section.

9.2.1.1.2 Equivalent rectangular waveguide model

Kumar and coworkers[9] have shown that the birefringence of an elliptical-core fiber closely matches that of a rectangular-core waveguide having the same core area, aspect ratio, and core and cladding refractive indices. Thus, the birefringence of an elliptical-core fiber with major and minor axes as 2*a* and 2*b*, respectively, is approximately equal to that of a rectangular-core waveguide with major and minor axes given by

$$2a' = \sqrt{\pi}a \text{ and } 2b' = \sqrt{\pi}b, \tag{9.6}$$

with the same core and cladding refractive indices, n_1 and n_2, respectively. The propagation constants β_x and β_y for the two fundamental modes of this rectangular core waveguide can be obtained by using an accurate perturbation approach[10] and are given by

$$\beta_i^2 = (\beta_{1i}^2 + \beta_{2i}^2) + k_0^2(n_1^2 - n_2^2)P_i^2, \tag{9.7}$$

where i stands for x or y, and β_{1i} and β_{2i} are obtained by solving the two following simple transcendental equations:

$$\mu_1 \tan \mu_1 = \alpha_i \sqrt{V_1^2 - \mu_1^2}, \tag{9.8}$$

and

$$\mu_2 \tan \mu_2 = \frac{1}{\alpha_i}\frac{n_1^2}{n_2^2}\sqrt{V_2^2 - \mu_2^2}, \tag{9.9}$$

where $\alpha_i = \dfrac{n_1^2}{n_2^2}$ (or 1), for $i = x$ (or y),

$$\mu_1 = a'\sqrt{k_0^2\frac{n_1^2}{2} - \beta_{1i}^2}, \qquad \mu_2 = b'\sqrt{k_0^2\frac{n_1^2}{2} - \beta_{2i}^2},$$

$$V_1 = k_0 a'\sqrt{n_1^2 - n_2^2}, \qquad V_2 = k_0 b'\sqrt{n_1^2 - n_2^2} = \frac{b'}{a'}V_1,$$

and

$$P_i^2 = \left[1 + \frac{1}{\alpha_i^2}\left(\frac{V_1^2}{\mu_1^2} + 1\right)\left(\frac{2\mu_1 + \sin 2\mu_1}{1 + \cos 2\mu_1}\right)\right]^{-1} \times \left[1 + \alpha_i^2\frac{n_2^4}{n_1^4}\left(\frac{V_2^2}{\mu_2^2} + 1\right)\left(\frac{2\mu_2 + \sin 2\mu_2}{1 + \cos 2\mu_2}\right)\right]^{-1}. \tag{9.10}$$

The birefringence B_g can now be obtained by determining β_x and β_y corresponding to E_{11}^x and E_{11}^y modes, respectively, using Eqs. (9.7)–(9.10).

Figure 9.3 shows the variation of the normalized birefringence $\Delta\tilde{\beta} = B_g / k_0(\Delta n)^2$ as a function of the V parameter $\left(= k_0 b\sqrt{n_1^2 - n_2^2}\right)$ for elliptical-core fibers having $n_2 = 1.47$, $\Delta\left[= \left(n_1^2 - n_2^2\right)/2n_1^2\right] = 0.01$, and different core ellipticities.[9] The solid curves correspond to the numerical calculations of Dyott et al.,[4] and the dashed curves are obtained by using Eqs. (9.7)–(9.10). The vertical lines denote the first higher-order-mode cutoffs taken from Ref. 11. As can be seen from the figure, the two curves match very closely in the region of practical interest—namely, for V values corresponding to maximum birefringence and just before the first higher-order-mode cutoff.

As mentioned earlier, the core dimensions of such fibers, required for single-mode operation with good polarization-holding capacity, become extremely small. This problem can be partially overcome by introducing a depressed inner cladding in the refractive index profile,[12] i.e., by making the profile a W-type fiber, as the cutoff V value of the first higher-order mode is larger in W-type fibers.

9.2.1.1.3 Stress birefringence B_s

Since the thermal expansion coefficient of the core and cladding materials are generally different, an asymmetrical stress is frozen in elliptical-core fibers during their fabrication. This stress in turn introduces a linear birefringence B_s in the fiber through the elasto-optic effect. The amount of introduced birefringence is obtained by calculating the difference between the stresses generated at the fiber center, along the major and minor axis of the fiber, and is given by[13–16]

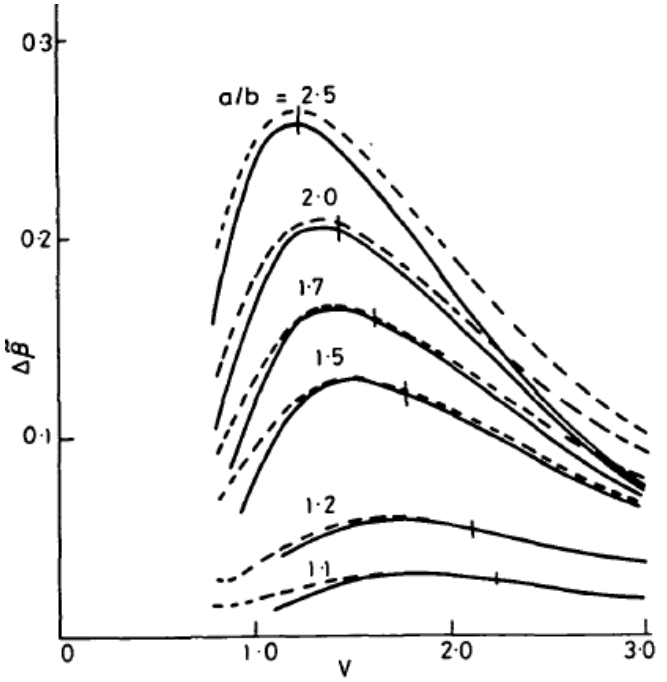


Figure 9.3 Variation of normalized $\Delta\tilde{\beta}$ as a function of the V parameter for elliptical core fibers, with n_2 = 1.47, $\Delta = 0.01$, and different core ellipticities. (Reprinted with permission from Ref. 9, © 1984, IEEE.)

$$B_s = k_0\, g(V) \frac{c_0}{1-\nu} \Delta\alpha\, \Delta T \frac{(a-b)}{(a+b)}, \tag{9.11}$$

where $c_0 = n_0^3 (p_{11} - p_{12})(1+\nu)/2$; n_0 is the average refractive index of the core and cladding; p_{11} and p_{12} are the components of the photoelastic tensor; ν is the Poisson's ratio of the fiber material; $\Delta\alpha$ is the difference between the thermal expansion coefficients of the core and cladding materials; $\Delta T = T_r - T_s$, with T_r and T_s being the room temperature and the softening temperature, respectively, of the fiber material; and $g(V)$ represents the fractional modal power in the core of the fiber. The factor $g(V)$ gives the wavelength dependence of B_s/k_0, as the fractional modal power in the core is strongly dependent on the wavelength. Approximating the elliptical core of the fiber by a circular core of radius $d = (a + b)/2$, $g(V)$ is approximately given by

$$g(V) \approx W^2 / V^2, \tag{9.12}$$

where $W = k_0 d \sqrt{n_{\text{eff}}^2 - n_2^2}$, $V = k_0 d \sqrt{n_1^2 - n_2^2}$, and n_{eff} is the average effective refractive index of the two fundamental modes.

9.2.1.2 Side-pit and side-tunnel fibers

Side-pit and side-tunnel fibers were proposed by Okoshi and coworkers.[17–19] Figure 9.4 shows the refractive index distribution in a transverse cross section of such fibers. In these fibers, as in the case for elliptical-core fibers, birefringence is obtained by introducing the geometrical anisotropy in the core. As shown in the figure, the fiber core consists of two pits of refractive index n_p (< cladding index n_2) on each side of the central core. Thus, the fiber has a W-type index profile along the x axis and a step-index profile along the y axis. If the two index pits are hollow (i.e., $n_p = 1$), the fiber is known as a *side-tunnel fiber*. Simple methods to obtain the birefringence in such fibers are either to use the effective index method[20] or to approximate it by a rectangular-core waveguide and then use a perturbation approach,[21] similar to the method used for elliptical core fibers.

When side-pit fibers were first proposed,[17] it was predicted that the birefringence of these fibers should be quite large, $\Delta n_{eff} \sim 1 - 2.5 \times 10^{-4}$. Later, it was also predicted[19] that in a side-tunnel fiber with optimized channels, even larger birefringence, $\Delta n_{eff} \sim 4.5 \times 10^{-4}$, can be obtained. However, in practical side-pit and side-tunnel fibers, the birefringence is found to be $\sim 5 \times 10^{-5}$ and $\sim 7 \times 10^{-5}$, respectively, which is of the same order as that of the elliptical-core fibers.

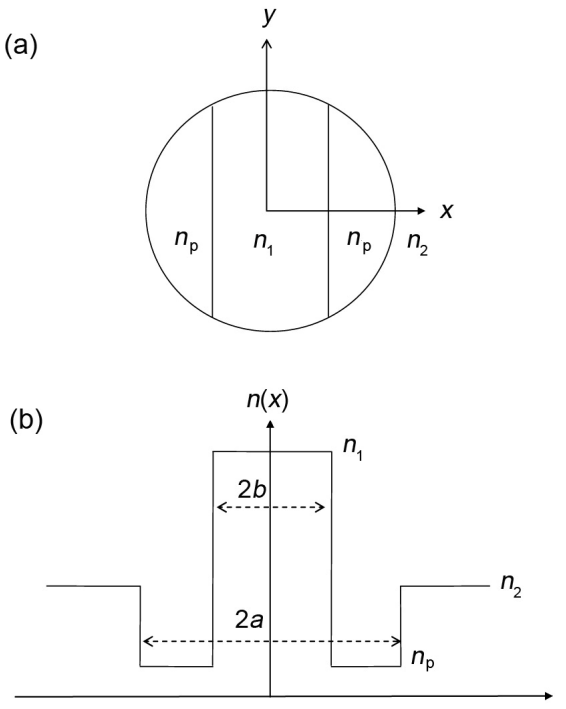


Figure 9.4 (a) Transverse cross section and (b) refractive index distribution along the x axis of a side-pit fiber.

9.2.1.3 Stress-induced fibers

A more effective method of introducing high birefringence in optical fibers is to introduce an asymmetric stress with twofold symmetry into the core of the fiber. The stress changes the refractive indices of the core (due to the photoelastic effect) seen by the two modes polarized along the principle axes of the fiber and generates birefringence. The required stress is obtained by introducing two identical stress-applying parts (SAPs) centered on a diameter, one on each side of the core.

The SAPs have a thermal expansion coefficient α_3 that differs from that of the cladding materials α_2, causing an asymmetrical stress to be applied on the fiber's core after it is drawn from the preform and cooled. The most commonly used SAPs are bow-tie shaped[22] and circularly shaped.[23–25] The corresponding fibers are known as bow-tie and polarization-maintaining and absorption-reducing (PANDA) fibers, respectively. Cross sections of these types of fibers are shown in Fig. 9.5. In the following section, we discuss their polarization characteristics and design considerations.

9.2.1.3.1 Modal birefringence

The asymmetrical stress in the fiber changes the refractive indices seen by the modes polarized along the principle axes (for instance, x and y) of the fiber. The effective refractive indices (n_x and n_y) seen by the x- and y-polarized modes, respectively, can be represented by the following equations:

$$\begin{aligned} n_x &= n_{x0} + c_1\sigma_x + c_2\sigma_y + c_2\sigma_z, \\ n_y &= n_{y0} + c_2\sigma_x + c_1\sigma_y + c_2\sigma_z, \end{aligned} \tag{9.13}$$

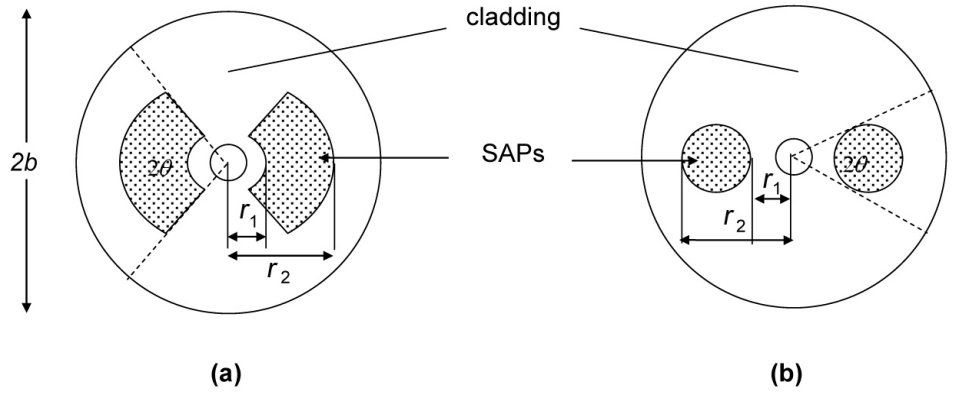


Figure 9.5 Transverse cross sections of (a) bow-tie and (b) PANDA stress-induced fibers.

where σ_x, σ_y, and σ_z are the stresses generated along the x, y, and z axis, respectively; c_1 and c_2 are the photoelastic coefficients; and n_{x0} and n_{y0} are the effective refractive indices of the x- and y-polarized modes in the absence of any stress. Thus, the modal birefringence B is given by

$$\begin{aligned} B &= k_0(n_x - n_y) \\ &= k_0[(n_{x0} - n_{y0}) + (c_1 - c_2)(\sigma_x - \sigma_y)] \\ &= (B_g + B_s). \end{aligned} \tag{9.14}$$

The term B_g naturally represents the geometrical (or shape) birefringence, while the term B_s represents the stress-induced birefringence. In the case of a circular-core fiber, the geometrical birefringence is negligibly small, and hence the modal birefringence is mainly determined only by the second term, which can be obtained by calculating the distribution of the difference of stresses σ_x and σ_y in and around the core.

The stress distribution due to SAPs in these fibers is quite cumbersome to obtain. Okamoto and coworkers[26] have provided a finite-element method to obtain this complicated stress distribution, while Varnham et al.[16] and Chu and Sammut[27] have provided analytical methods for the calculations of thermal stress distribution and hence the birefringence for such fibers. It is shown that in the region close to the center of the fiber, the stress-induced birefringence in the bow-tie and PANDA fibers are given by the following equations:

For bow-tie fibers,[27]

$$B \sim \frac{k_0}{(1-\nu)}(\alpha_2 - \alpha_3)\Delta T \frac{c_0}{\pi}\left[2\ln\left(\frac{r_2}{r_1}\right) - \frac{3}{2b^4}\left(r_2^4 - r_1^4\right)\right]\sin 2\theta\,, \tag{9.15}$$

where r_1 and r_2 represent the inner and the outer core radii, respectively, of the bow-tie sectors, 2θ represents the projection angles of each sector at the fiber center, and b represents the outer radius of the fiber (Fig. 9.5). Equation (9.15) shows that B is maximized for $2\theta = \pi/2$ (the optimum sector angle). Further, B increases as r_1 decreases, which means that the SAPs should be as close to the core as possible. The birefringence also increases by increasing r_2; however, it reaches a maximum for $r_2 = 0.76b$, beyond which if r_2 increases, the birefringence begins to decrease.[16]

For PANDA fibers,[27]

$$B \sim \frac{k_0 2c_0}{(1-\nu)}(\alpha_2 - \alpha_3)\Delta T\left(\frac{d_1}{d_2}\right)^2\left[1 - 3\left(\frac{d_2}{b}\right)^4\right], \tag{9.16}$$

where $d_1[=(r_2 - r_1)/2]$ represents the radius of the SAPs, and $d_2[=(r_1 + r_2)/2]$ represents the distance between the fiber center and the center of either of the SAPs, r_1 and r_2 being the distances of the closest and farthest edges of the SAPs from the fiber center. Equation (9.16) shows that B can be increased by decreasing d_2, again indicating that SAPs should be as close to the core as possible. As in the case of the bow-tie fiber, B can be increased by increasing r_2 (or d_1) up to an optimum value. The optimum value of r_2 in this case is given by[27]

$$\frac{r_2}{b} = r_0\left[1+\frac{2R_1^{1/4}r_0^{3/4} - R_1 - r_0}{(R_1 + 5r_0)/4}\right], \tag{9.17}$$

where $R_1 = r_1/b$, and $r_0 = (1/3)^{1/4} = 0.76$ is the optimum value of r_2/b for bow-tie fibers.

As discussed earlier, placing the SAPs closer to the core improves the birefringence of these fibers; however, they cannot be placed arbitrarily close to the core, since this increases the loss of the fiber, as SAPs are doped with materials other than silica. Stress-induced fibers are superior to any other type of birefringent fiber in terms of loss and polarization-holding capacity. Fibers with very low loss and low cross talk have been reported using optimum design parameters. For example, Sasaki et al.[28] have reported the fabrication of a PANDA fiber having 0.22-dB/km loss with birefringence $B/k_0 = 3.3 \times 10^{-4}$ and cross talk of −27 dB in a 5-km length.

9.2.1.4 Circularly birefringent fibers

Both the bow-tie and the PANDA-type fibers are linearly birefringent fibers, because the two orthogonally polarized modes of such fibers are linearly polarized. However, it is possible to introduce circular birefringence in the fiber so that the two orthogonally polarized modes of the fiber are right- and left-circularly polarized. The simplest way to achieve circular birefringence in a fiber is to twist it. If a fiber is twisted, it becomes circularly birefringent.[29] It is shown theoretically that if τ represents the twist rate per unit length, then the plane of polarization of a linearly polarized light is rotated in the same direction with a rate of $g\tau/2$, where $g \sim 0.16$ for silica fibers.[30] This means that a phase difference is generated between the two circularly (right and left) polarized modes, and the corresponding birefringence is given by

$$\Delta\beta = g\tau. \tag{9.18}$$

However, the birefringence generated in this manner is very small, as g is small. It is very difficult to obtain beat lengths less than about 10 cm using this method, as the fiber will break at higher twist rates. Furthermore, a twisted fiber is very difficult to handle.

9.2.2 Single-polarization single-mode (SPSM) fibers

In Section 9.2.1, we discussed high-birefringence fibers in which the SOP of the light beam is maintained by reducing the coupling between the two fundamental modes. Another class of polarization-maintaining fibers supports only one polarized mode. Such fibers are known as single-polarization single-mode (SPSM) fibers, which are basically in-line fiber optic polarizers. In the literature, a large number of schemes have been reported to achieve SPSM fibers, a few of which are discussed in the following sections.

9.2.2.1 Fibers with different cutoffs for the two modes

In this type of fiber, the two fundamental modes have different cutoff wavelengths. Thus, at wavelengths lying between the two cutoffs, one of the modes is not guided and the fiber behaves as an SPSM fiber. Side-pit and side-tunnel fibers, discussed in Sec. 9.2.1.2, are one example of such fibers. In these fibers, the cutoff frequencies of the *x*- and *y*-polarized modes (V_{cx} and V_{cy}, for instance) are different, and thus over a certain wavelength range, these fibers support only one polarized mode. Kaul and coworkers[31] have shown that the cutoff frequencies V_{cx} and V_{cy} in side-pit and side-tunnel fibers are nearly the same as those of a W-type planar waveguide whose refractive index profile is the same as that of the given fiber along the *x* axis [as shown in Fig. 9.4(b)]. Hence, the cutoff of the *x*- (*y*-) polarized mode of the fiber should be determined by the cutoff of the fundamental TM (TE) mode of the W-type planar waveguide [shown in Fig. 9.4(b)], which can be obtained by solving the simple transcendental equation[31]

$$\tan\left(V_c \frac{b}{a}\right) + \eta\gamma \tanh\left[\gamma V_c \left(\frac{b}{a} - 1\right)\right] = 0, \tag{9.19}$$

where

$$\eta = \frac{n_1^2}{n_2^2} \quad \text{for TM modes,}$$

$$= 1 \quad \text{for TE modes,}$$

$$\gamma = \left(\frac{n_1^2 - n_p^2}{n_1^2 - n_2^2}\right)^{1/2},$$

and n_1, n_2, and n_p are the refractive indices of the central core, cladding, and pit regions, respectively, and V_c represents the value of the normalized frequency $V\left(= k_0 a\sqrt{n_1^2 - n_2^2}\right)$ at the cutoff. This gives a simple method for calculating the relative single-polarization bandwidth *S*, defined as

$$S = 2\left[\frac{V_{cx} - V_{cy}}{V_{cx} + V_{cy}}\right], \tag{9.20}$$

for side-pit and side-tunnel fibers. The predicted single-polarization bandwidth of side-pit fibers[32] is ~2.5% and for side-tunnel fibers is ~7%.

9.2.2.2 Single-guided polarized mode

Eickhoff[13] has proposed a way to make one of the polarized modes unguided through stress effects. He has shown that for fibers with elliptical core or elliptical inner cladding, the stress-induced refractive index profiles for the *x*- and *y*-polarized modes are different. Furthermore, the direction of the fast axis of the stress-induced birefringence depends on the photoelastic constants and the thermal expansion coefficients of the fiber materials [see Eq. (9.11)]. Thus, by appropriately selecting the fiber material, it should be possible for the stress-induced refractive index profile to cancel the dopant-induced refractive index profile (which should be the same for both the *x*- and *y*-polarized modes) for one of the polarized modes, so that the fiber does not act as a waveguide for this mode.

Two such possibilities are shown schematically in Fig. 9.6. In the first case [Fig. 9.6(a)], the dopant-induced refractive index profile is a negative index step, while the stress induces the positive index steps such that for the *y*-polarized mode, the two profiles cancel each other, while the stress-induced positive index step for the *x*-polarized mode is not exactly cancelled. As a result, the *y*-polarized mode is not guided. Similarly, the *x*-polarized mode can also be made unguided, as shown in Fig. 9.6(b). Snyder and Ruhl[33] have proposed the use of birefringent materials to obtain similar refractive index profiles for *x*- and *y*-polarized modes.

9.2.2.3 High differential leakage loss between the two modes

Single-mode single-polarization fibers can also be realized if the two polarized modes have a large difference between their losses. In the following, we discuss two different methods to obtain the differential leakage loss between the two modes.

9.2.2.3.1 By bending a Hi-Bi fiber

One way to obtain the differential leakage loss between the two modes is by bending a Hi-Bi fiber. Varnham and coworkers[34] used a bow-tie fiber, while Okamoto and coworkers[35] used a PANDA fiber to obtain the single-polarization operation through bending. Figure 9.7 shows the spectral attenuation plot of a bent bow-tie fiber.[34] It is clear from this figure that the two polarized modes have a differential loss around λ_0 = 800 nm and that the fiber behaves like a single-polarization fiber. Okamoto and coworkers[35] used a bent flat cladding PANDA fiber to achieve SPSM operation. A polarization extinction ratio of more than 45 dB was reported[35] at λ_0 = 1.3 μm.

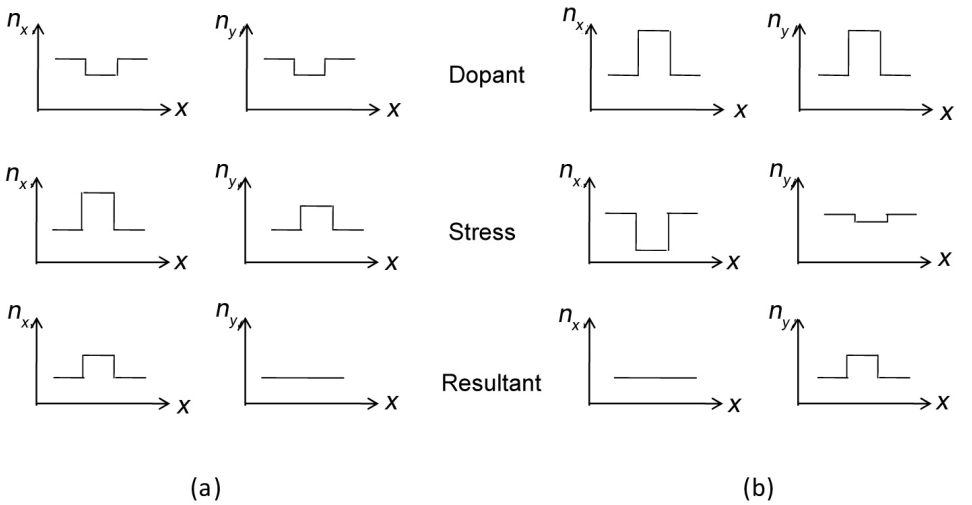


Figure 9.6 Dopant-induced, stress-induced, and the resultant refractive index profiles of possible schemes to make (a) *y*-polarized and (b) *x*-polarized unguided modes for achieving SPSM fibers (adapted from Ref. 13).

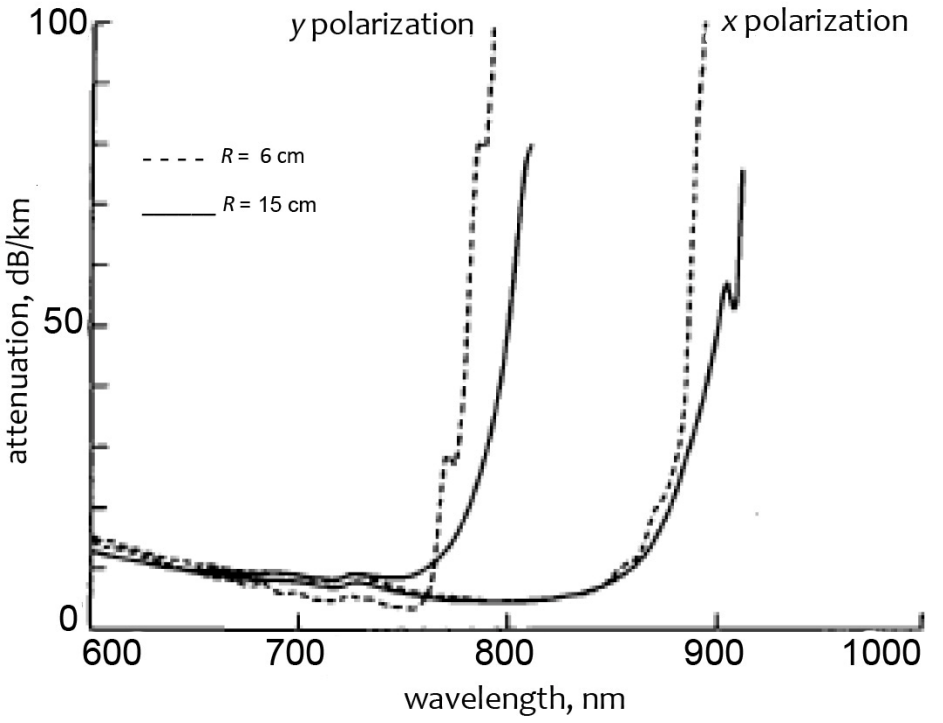


Figure 9.7 Spectral attenuation of *x*- and *y*-polarized modes in a bow-tie fiber for two different bent radii *R*. (Reprinted with permission from Ref. 34, © 1983, IEEE.)

9.2.2.3.2 Using fibers made with birefringent material

Snyder and Ruhl[33] proposed the use of fibers composed of highly birefringent materials to obtain differential leakage loss between the two modes. Consider a fiber whose core and cladding are both made with highly birefringent material so that the refractive index profile seen by the *x*- and *y*-polarized modes is like the one shown in Fig. 9.8. Here, n_{1x} (n_{1y}) and n_{2x} (n_{2y}) represent the core and the cladding indices for *x*- (*y*-) polarized modes. It is well known that the effective refractive index for the *x*- (*y*-) polarized mode would then lie between n_{1x} (or n_{1y}) and n_{2x} (or n_{2y}). It can easily be shown that for *V* parameters smaller than a certain value, the value of (β_y / k_0) can be made smaller than n_{2x}. Since the *y*-polarized mode also has a tiny *x* component (as the modes of a fiber are not exactly linearly polarized; see Fig. 8.7), this mode becomes leaky, as the tiny *x* component behaves as if it is traveling in a hollow waveguide. Theoretical calculations made by Synder and Ruhl[33] show that leakage losses of the *y*-polarized mode of the order of 1000 dB/km should be possible.

9.3 Applications of Birefringence in Optical Fibers

Birefringence has many applications in optical fiber technologies. One requirement in many of these applications is that the polarization of the output light is stable and well defined. In such cases, it is preferable to use Hi-Bi fibers over conventional single-mode fibers.

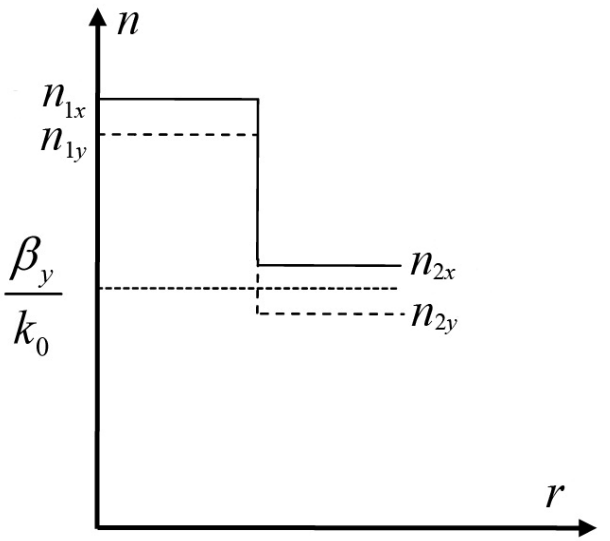


Figure 9.8 Schematic refractive index profile for *x*- and *y*-polarized modes of SPSM fibers proposed by Snyder and Ruhl (adapted from Ref. 33).

9.3.1 Applications requiring a stable polarization state

9.3.1.1 Coupling to integrated optical circuits

Integrated optical waveguides are polarization sensitive, as the fractional power contained in the core of this type of waveguide is different for TE and TM modes. As a result, the polarization direction of light coming out of an optical fiber should be stable so that the power coupled to an integrated optical receiver is also stable. Fiber links connected to these devices must therefore preserve a well-defined SOP. The unstable SOP at the output end of an ordinary single-mode fiber causes fading of the received signal.

9.3.1.2 Interferometric sensors

Fiber interferometric sensors also require light of a well-defined SOP for efficient operation, since interferometric fringes produced by two beams are due to their components being polarized in the same direction. If the polarization states of the two interfering beams are orthogonal to each other, no interference can occur. With an ordinary single-mode fiber, the output SOP can change, due to environmental factors such as bending and twisting. These fluctuations in the SOP are not related to the physical parameter that is being sensed, and they produce polarization noise and signal fading. This problem can be avoided by using a Hi-Bi fiber with light launched along one of the principal axes of the fiber, since environmental factors do not affect the polarization state of the propagating light in such fibers.

In a fiber optic gyroscope, the use of a polarization-maintaining fiber is preferred over the conventional single-mode fiber. A fiber optic gyroscope measures rotation by detecting the phase difference between the beam that propagates clockwise through the fiber coil and the beam that propagates counterclockwise through the fiber coil. One of the requirements for the proper functioning of this sensor is that both paths are identical, which can be satisfied only if both beams travel through the same SOP. If there is a difference in optical paths for the two beams, the output signal can drift, giving errors in the measured rotation. Use of a Hi-Bi fiber ensures that both of the beams travel through the same SOP and have identical paths.

9.3.1.3 Coherent communication systems

One major application of polarization-maintaining fibers is in the field of coherent communications. In coherent communication systems, a light signal transmitted through a single-mode fiber is mixed with light from a local oscillator at the receiver side to produce a beat signal. It can be shown that the signal-to-noise ratio is much improved in this detection scheme, as compared to direct detection. However, polarization states of the transmitted signal and the local oscillator signal should constantly match in order to obtain the highest sensitivity and stability in detection. This also requires the use of PMFs.

9.3.2 In-line fiber optic devices and components

Optical fibers were once thought of only as a transmission medium. It was proposed that any processing of the optical signal would take place outside of the fiber. However, in the recent past, a large number of fiber optic in-line components have been researched and developed in which the guided light is processed inside the fiber itself. Before the development of PMFs, only a few components such as polarization-independent couplers and star couplers had been developed. However, with the development of polarization-maintaining fibers, a wide range of fiber devices such as polarization-maintaining couplers, filters, multiplexers, and demultiplexers have been reported.[36,37] In the remainder of Section 9.3 we consider various fiber optic devices and components that use either polarization-maintaining fibers or the controlled birefringence introduced in normal single-mode fibers.

9.3.2.1 Two-mode elliptical-core fiber sensors

Two-mode elliptical-core fiber (TMEF) sensors are based on the differential phase modulation between the LP_{01} and $\mathrm{LP}_{11}^{\mathrm{even}}$ modes of the fiber that is due to the physical parameter being measured. The working principle of such a sensor is described in Fig. 9.9. The field at the output end of a TMEF is the superposition of the fields due to the LP_{01} and $\mathrm{LP}_{11}^{\mathrm{even}}$ modes. If both of the modes are excited with equal power at the input end of the fiber, one obtains a two-lobe pattern in its near-field region (at the fiber's output end) as well as in the far-field region, as shown in Fig. 9.9(a). The intensity of either of the lobes varies periodically as the phase difference between the two modes at the output end changes. The phase difference $\Delta\phi$ is given by

$$\Delta\phi = (\beta_{01} - \beta_{11}^{\mathrm{even}})L \,. \tag{9.21}$$

In the preceding equation, β_{01} and $\beta_{11}^{\mathrm{even}}$ are the propagation constants of the LP_{01} and $\mathrm{LP}_{11}^{\mathrm{even}}$ modes, respectively, and L is the length of the fiber. The relative intensities of the two lobes in the near-field and far-field regions for different values of $\Delta\phi$ are shown schematically in Figs. 9.9(b) and (c), respectively. Thus, by monitoring the intensity of any of the lobes in the near-field or far-field region, one can measure the change in the phase difference between the two modes and hence the physical parameter causing it. Several TMEF sensors that measure parameters such as axial strain, temperature, and acceleration have been reported in the literature.[38–41]

Using TMEFs, many in-line fiber optic components such as acousto-optic frequency shifters,[42,43] modal couplers,[44] dispersion compensators,[45,46] and mode splitters[47] have been reported in the literature. In the next section, we briefly discuss a dispersion compensator using TMEFs.

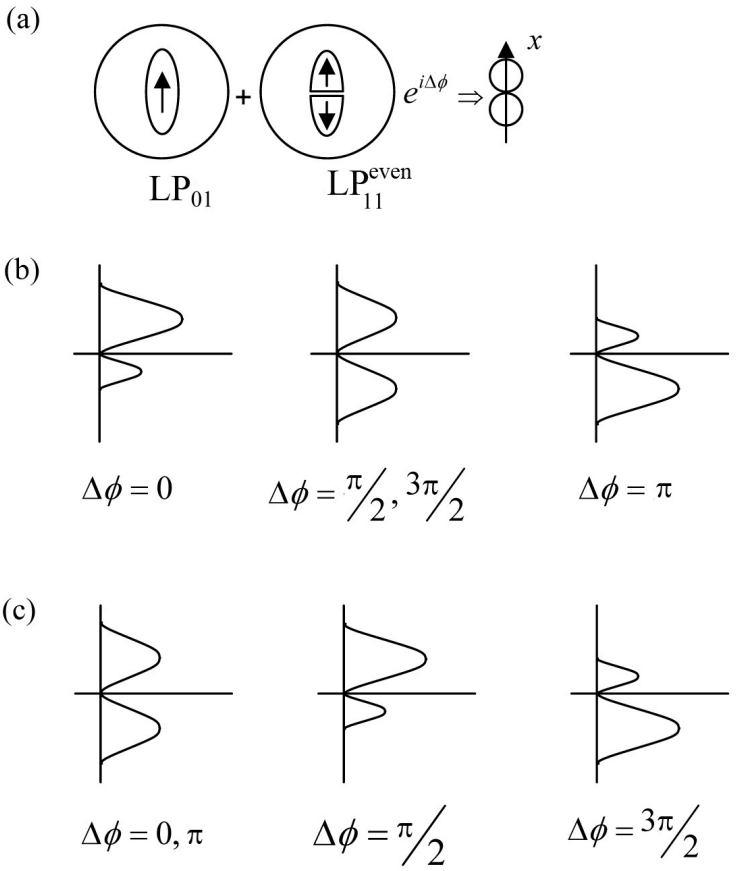


Figure 9.9 (a) Schematic representation of LP_{01} and LP_{11}^{even} modes of a two-mode elliptical-core fiber having a phase difference of $\Delta\phi$. A combination of the two modes gives a two-lobe pattern; the relative intensities of the two lobes for different values of $\Delta\phi$ as obtained in (b) the near-field and (c) the far-field regions are shown.

9.3.2.2 Dispersion compensator using a two-mode elliptical-core fiber

Already installed fiber-optic communication links employ standard single-mode fibers (SSMFs) that have a positive dispersion coefficient (D ~ 17 ps/km·nm) at a wavelength of 1550 nm. This dispersion results in a temporal broadening of the propagating optical pulses and can lead to interchannel interference, which ultimately leads to a degradation in the performance of the link. Thus, dispersion compensators are required, which should have a high negative dispersion coefficient in order to equalize (or compensate for) the positive dispersion accumulated over a certain length of the SSMF (which is typically 50 km).

Poole et al.[45] proposed an equalization scheme using two-mode fibers (TMFs) that uses the fact that the LP_{11} mode of the TMF has a large negative dispersion close to its cutoff. In this scheme, the LP_{01} mode propagating in the

SSMF is converted to the LP_{11} mode of a TMF section (acting as the dispersion compensator). At the end of this section, the LP_{11} mode is again converted back to the LP_{01} mode of the SSMF. Conversion between the LP_{01} and LP_{11} modes of the SSMF and TMF, respectively, is achieved by a spatial mode coupler using periodic microbending. In another work, Poole et al.[46] used a TMEF operated in the near-cutoff region for the LP_{11} mode with a measured dispersion of –548 ps/km·nm for the LP_{11} mode. Using a double-pass configuration, complete compensation of the dispersion in a 50-km-long SSMF at 1530 nm was reported[46] with just 0.75 km length of the TMEF. Figure 9.10 gives their experimental results, showing the input and output bit patterns with and without using the dispersion compensator.

9.3.2.3 Fiber optic polarization beamsplitter

A *fiber optic polarization beamsplitter* is a fiber version of the Wollaston prism of bulk optics. Such a device is extremely useful in the analysis of polarized light and in the realization of a circulator and an optical time-domain reflectometer (OTDR). The device can be realized through appropriate directional coupling between two high-birefringence fibers such as PANDA fibers.

It may be mentioned here that if we bring the cores of two SMFs close enough such that the evanescent fields of the modes supported by the individual fibers begin to overlap, then a periodic exchange of energy occurs between the two fibers. Such a structure is known as a *directional coupler*.[48] The strength of coupling between the two waveguides is given in terms of a parameter known as the *coupling length* L_c, which is defined as

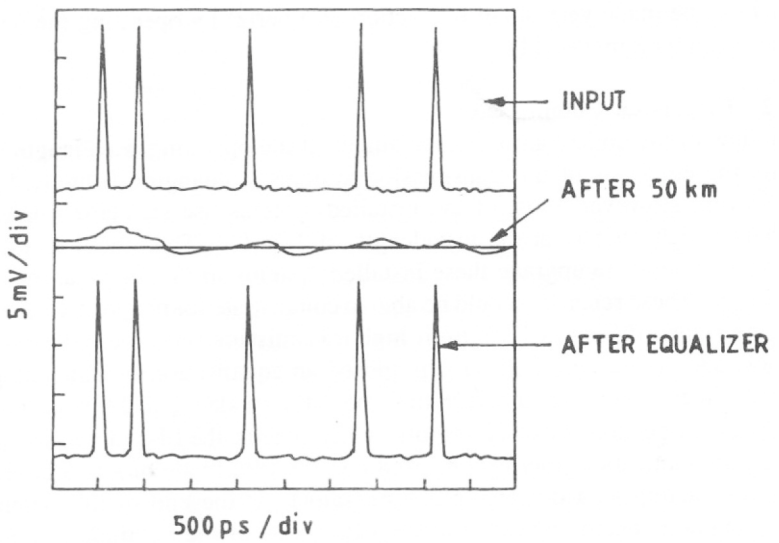


Figure 9.10 Bit patterns measured at the input and after 50 km of a SMF with and without dispersion compensation. (Reprinted with permission from Ref. 46, © 1993, IEEE.)

$$L_c = \frac{2\pi}{\beta_s - \beta_a}, \tag{9.22}$$

where β_s and β_a represent the propagation constants of the symmetric (even) and the antisymmetric (odd) modes (also known as the *supermodes*) of the coupled structure. It can be shown that if the power is coupled into one of the fibers at the input, then at the output, all of the power will be coming out of the coupled fiber when the interaction length L between the two fibers is given by

$$L = (2m+1)L_c, \tag{9.23}$$

and will be coming out of the same fiber when

$$L = 2mL_c, \tag{9.24}$$

where m is an integer. In the case of a Hi-Bi-fiber directional coupler, if the fiber axes are aligned, the coupling lengths for the two orthogonal polarization modes becomes significantly different. It is thus possible to make a directional coupler using Hi-Bi fibers such that the interaction length L satisfies the following condition:

$$L = 2mL_{cx} = (2n+1)L_{cy}, \tag{9.25a}$$

or

$$L = (2m+1)L_{cx} = 2nL_{cy}, \tag{9.25b}$$

where m, n = 1, 2, 3, …. In such a case, the x and y components of the input SOP will be split at the output ports. Yokohama and coworkers[49] reported such a directional coupler using two PANDA fibers, as seen in Fig. 9.11. The figure shows that the input SOP is linear and oriented at an angle θ with the x axis and that the x- and y-polarization components are split at the output.

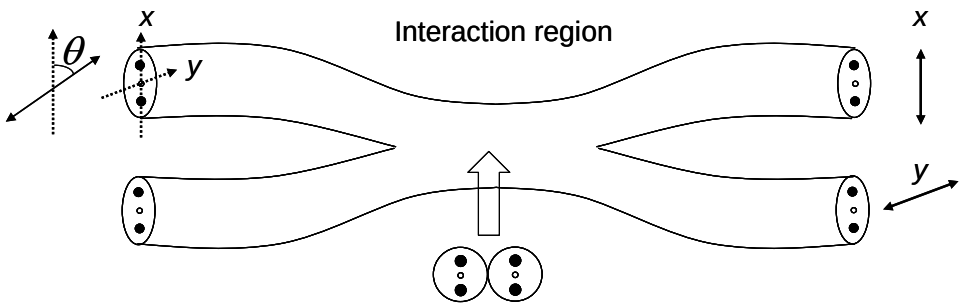


Figure 9.11 Schematic of a directional coupler using two PANDA fibers with aligned polarization axes acting as a polarization beamsplitter.

The polarization beamsplitter in Ref. 49 had an interaction length of 20 mm and was shown to act as a polarization beamsplitter at wavelengths of 1.282 μm, 1.346 μm, and 1.402 μm.

9.3.3 Fiber optic devices using controlled birefringence in SSMFs

The fiber optic devices discussed in Section 9.3.2 are based on Hi-Bi fibers. A large number of in-line fiber optic polarization components and devices such as fiber polarizers, all-fiber wave plates (e.g., QWPs and HWPs), and current sensors use a controlled birefringence introduced in normal single-mode fibers. In the remainder of Section 9.3.3, we discuss some of these fiber optic devices.

9.3.3.1 Zero-birefringence optical fiber holder

In many applications such as interferometric and polarimetric sensors, the fiber needs to be held without introducing extra birefringence in the fiber. However, in order to fix the fiber securely in its intended position, lateral forces need to be applied to the fiber; these forces produce birefringence in the fiber, as discussed in Section 8.6.2.2. Such a birefringence is highly undesirable, as it tends to vary with time, e.g., with changing ambient temperature. A possible solution to this problem is to clamp the fiber in a V groove with a groove angle of 60 deg.

Figure 9.12 shows the cross-sectional view of a bare fiber of outer radius r pressed in a V groove of angle 2γ. If a force F per unit length is applied on the fiber in the y direction, two reactional forces F_1 and F_2 at the points of contact result. Kumar and Ulrich[50] have studied both theoretically and experimentally the birefringence produced in this way. In an ideal case of zero friction in the groove, F_1 and F_2 are perpendicular to the groove walls, as shown in Fig. 9.12, and the birefringence is given by[50]

$$B_s = \beta_x - \beta_y = -\frac{hk_0}{2}\left(1 - \frac{\cos 2\gamma}{\sin \gamma}\right)\frac{F}{r}. \tag{9.26}$$

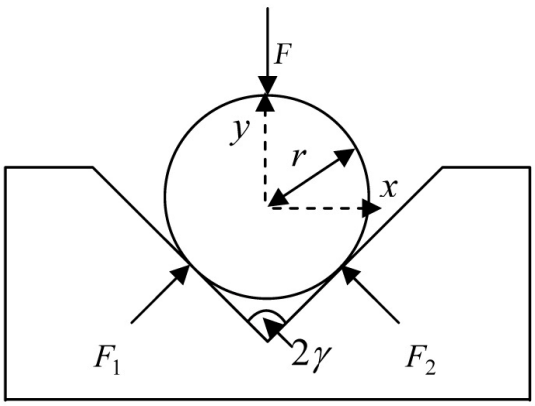


Figure 9.12 Fiber pressed in a V groove.

In practice, however, the friction in the groove is not zero, and the preceding equation takes the form[50]

$$B_s = -\frac{hk_0}{2}\left(1-\frac{\cos 2\gamma}{\mu\cos\gamma+\sin\gamma}\right)\frac{F}{r}, \tag{9.27}$$

where μ is the coefficient of friction between the fiber and the groove's walls, and h is given by Eq. (8.81). The preceding equation reduces to Eq. (9.26) in the limiting case of zero friction, i.e., $\mu = 0$. In the other limiting case of complete friction, forces F_1 and F_2 are in exactly the opposite direction to the applied force F, and the corresponding value of the coefficient of the static friction is[50]

$$\mu = \frac{1}{\tan\gamma}, \tag{9.28}$$

and the induced birefringence is given by

$$B_s = -\frac{hk_0}{2}\left(1-\cos 2\gamma\sin\gamma\right)\frac{F}{r}. \tag{9.29}$$

Figure 9.13 shows the normalized birefringence as a function of groove angle for zero (solid curve) and complete friction (dashed curve), which are obtained using Eqs. (9.26) and (9.29), respectively. In the zero-friction case, the birefringence becomes zero and changes sign at $2\gamma = 60$ deg. However, in the case of complete friction, the birefringence would never be zero but has a minimum value at $2\gamma \approx 48.2$ deg.

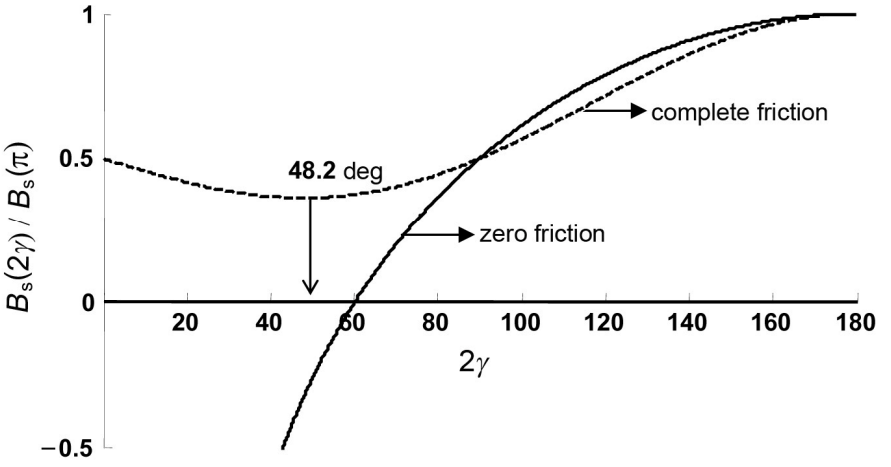


Figure 9.13 Normalized birefringence introduced in a fiber when pressed in a V groove. Solid and dashed curves correspond to Eqs. (9.26) and (9.29), respectively.

In practice, the introduced birefringence will lie somewhere between the two cases, depending on the quality of the groove. Maystre and Bertholds[51] have reported birefringence measurements by preparing different V grooves with good surface quality. Their studies show that V grooves made with hard steel produce zero birefringence at an angle of $2\gamma \approx 55$ deg.

9.3.3.2 In-line fiber optic wave plates

As previously mentioned, bending in the optical fiber leads to a linear birefringence, i.e., the two eigen SOPs of the system are linearly polarized. The fast SOP is oriented in the plane of the bend (for instance, the *y* axis), and the slow SOP is oriented perpendicular to the bend (for instance, the *x* axis). In this case, the amount of birefringence can be controlled by varying the bend radius and the number of bends or loops in the fiber. It can easily be shown using Eq. (8.84) that for a typical silica fiber, the phase difference accumulated between the two eigen SOPs in *N* fiber loops each of radius *R* (see Fig. 9.14) is given by

$$\Delta\phi = k_0 \times 0.136\left(\frac{r}{R}\right)^2 \times (2\pi RN) = 0.136 \times \frac{4\pi^2 r^2}{\lambda R} N .$$

Thus, a single fiber loop ($N = 1$) would act as a QWP (producing $\Delta\phi = \pi/2$) if the bend radius is given by

$$R = 0.136 \times \frac{8\pi r^2}{\lambda} . \tag{9.30}$$

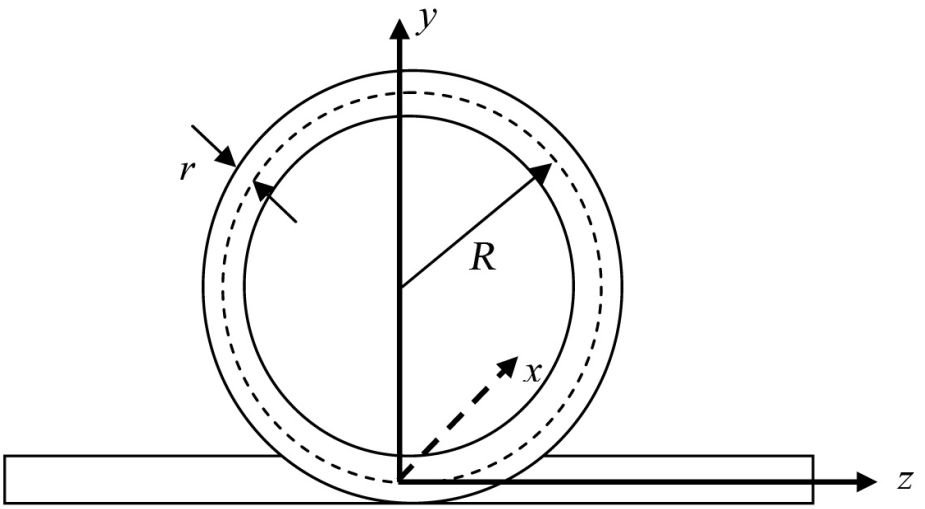


Figure 9.14 A fiber bent in the form of a loop.

Similarly, two loops ($N = 2$) instead of one with the same bending radius would give rise to a phase difference of π and would act as an HWP. Substituting $r = 62.5\times10^{-6}$ m and $\lambda = 1.310$ μm, one obtains $R \sim 1$ cm. These wave plates can be used either independently or in a suitable combination to transform the SOP of incident light from one state to another, as in a polarization controller, discussed in the next section.

9.3.3.3 All-fiber polarization controller

A proper sequence of three wave plates—a QWP, followed by an HWP, and then followed by another QWP with proper orientations—can convert the SOP of the input light to any desired output SOP. The resulting device is known as a *polarization controller*[52] (see Fig. 9.15). The device works in three steps: (1) converting the given elliptic SOP to a linear SOP, (2) rotating the linear SOP to another appropriate linear SOP, and (3) imparting to this new linear SOP the required ellipticity to obtain the final desired SOP (see Section 7.4).

The advantage of this device is that it is in an all-fiber form. In many interferometric and polarimetric fiber sensors, maintaining control over the SOP of light propagating in the fiber is required. In those cases, this polarization controller is immensely useful, as it is simply spliced at two ends between a fiber link, so bulk-optic components, which are lossy as well as difficult to integrate with optical fibers, are not needed.

9.3.3.4 Fiber optic current sensor

As discussed in Section 8.6.2.5, a longitudinal (i.e., axial) magnetic field **H** produces a circular birefringence in the fiber via the Faraday effect. This birefringence is used to measure electric currents. A single-mode optical fiber is coiled around a current-carrying conductor, and the magnetic field lines generated by the current are parallel to the fiber axis and, through the Faraday effect, will rotate the direction of an incident linearly polarized light, given by

$$\theta = \frac{B_{\mathrm{H}} l}{2} = V_{\mathrm{d}} \int \mathbf{H} \bullet \mathbf{dl} = V_{\mathrm{d}} NI\,, \tag{9.31}$$

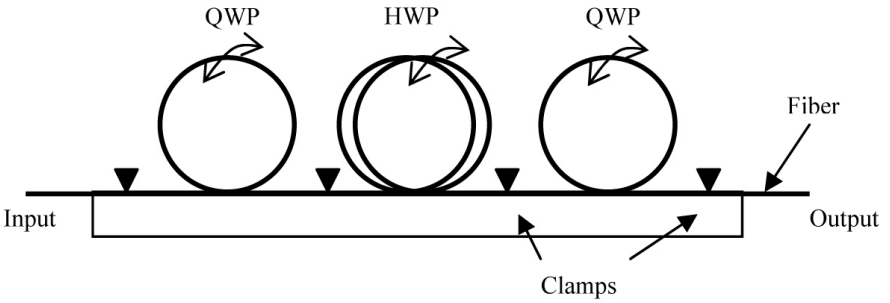


Figure 9.15 Schematic of an in-line fiber optic polarization controller.

where we have used Ampere's law $\int \mathbf{H} \cdot \mathbf{dl} = NI$, N being the number of turns in the fiber coil, and I representing the current flowing in the conductor.

Thus, by measuring the rotation of the polarization plane, one can measure the electric currents. A schematic of a fiber optic current sensor is shown in Fig. 9.16. Light from a linearly polarized laser is passed through an HWP so that its plane of polarization can be rotated. At the output end, the light beam will be linearly polarized (neglecting the bending-induced birefringence). When there is no current flowing in the conductor, the orientation of the output SOP is assumed to be along the x axis. The output is passed through a Wollaston prism rotated by 45 deg so that its two output components correspond to linear +45-deg and –45-deg components with respect to the x axis. If the output SOP makes an angle θ with the x axis, the intensities of these components I_{+45} and I_{-45} are given by

$$I_{+45} = I_0 \cos^2\left(\frac{\pi}{4} - \theta\right) \quad \text{and} \quad I_{-45} = I_0 \cos^2\left(\frac{\pi}{4} + \theta\right). \tag{9.32}$$

The angle θ can then be obtained by calculating

$$\frac{I_{+45} - I_{-45}}{I_{+45} + I_{-45}} = \frac{\cos^2\left(\frac{\pi}{4} - \theta\right) - \cos^2\left(\frac{\pi}{4} + \theta\right)}{\cos^2\left(\frac{\pi}{4} - \theta\right) + \cos^2\left(\frac{\pi}{4} + \theta\right)} = \sin 2\theta. \tag{9.33}$$

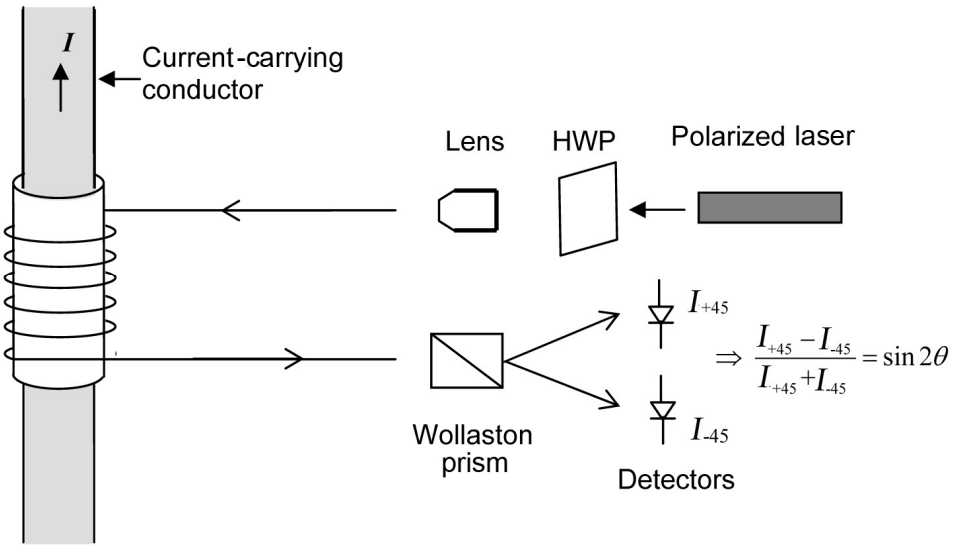


Figure 9.16 Schematic of a fiber optic current sensor.

This measurement technique is free from intensity fluctuations in the input light beam.

Example 9.1 Assuming that the fiber coil around the current-carrying conductor consists of 100 turns, then for a current of 1 A, the polarization plane of the incident light would be rotated by

$$\theta = 4.7\times10^{-6}\times100\times1$$
$$= 4.7\times10^{-4} \text{ rad} \approx 2.7\times10^{-2}\,\text{deg}.$$

9.3.3.5 In-line fiber optic polarizer

As discussed in Section 8.6.2.6, a metallic layer near the fiber core makes the TM mode highly lossy. This metallic layer is used to realize TE-pass in-line fiber optic polarizers. A small portion of SMF is fixed in a quartz block after removing its plastic cladding and is polished from one side up to very near its core. Such a structure is known as a *polished half-block*. We then deposit a layer of appropriate metal (gold, silver, or aluminum) on the polished surface, which acts as a TE-pass polarizer [Fig. 9.17(a)].

A more efficient design of a metal-clad polarizer uses a thin metallic film of appropriate thickness (< 10 nm) capped with an appropriate dielectric, as shown in Fig. 9.17(b). A thin metallic layer surrounded by dielectric media supports symmetric and antisymmetric combinations of surface plasma polaritons supported by opposite metal–dielectric interfaces.[53,54] The propagation constants of these super SPPs depend on the dielectric constants of the dielectric media on the two sides of the metal layer. Thus, if one selects the capping material appropriately, the effective index of the fiber mode will match the effective index of one of the super SPPs. In that case, the fiber TM mode will be coupled to that SPP and leaked out, resulting in an efficient TE-pass polarizer. Polarization extinction ratios as high as 70 dB have been reported, with TE-mode loss as low as 0.5 dB.[54]

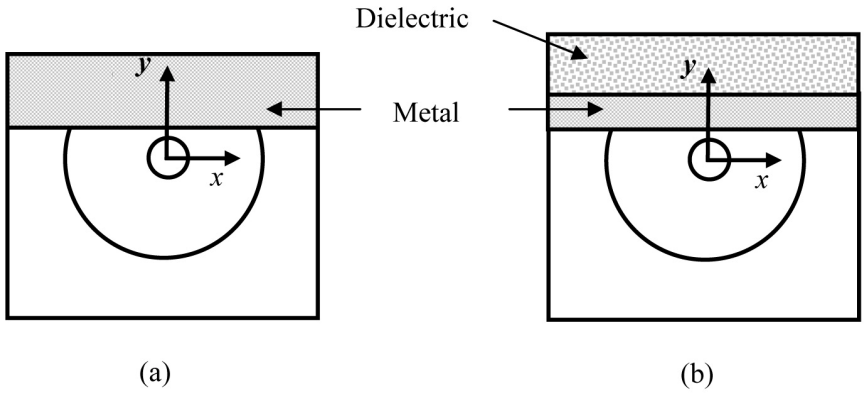


Figure 9.17 Cross-sectional view of a side-polished optical fiber with (a) a thick metal layer and (b) a thin metal layer capped with a dielectric.

In the recent past, a large number of highly sensitive fiber optic sensors have been reported using coupling between a fiber mode and SPPs. In the following section, we briefly discuss the working principal of such devices.

9.4 Surface Plasmon Polaritons and Devices

Surface plasmons (SPs) are coherent electron oscillations supported at a dielectric–metal interface having opposite (positive–negative) real parts of their dielectric constants. When SPs couple with photons, the resulting hybridized excitation is called a surface plasmon polariton (SPP), which propagates along the metal–dielectric interface until energy is lost via absorption in the metal. SPPs have their field maxima near the interface, and they decay exponentially beyond the interface in both mediums, as shown in Fig. 9.18. Furthermore, their complex propagation constant with a large imaginary part makes them highly lossy. The existence of SPs was first predicted in 1957 by Ritchie,[55] and since then, SSPs have been extensively studied for their possible applications in sensing and a number of other applications such as optoelectronic devices.

Following a simple analysis for the surface mode supported at a metal–dielectric interface, it can be shown that SPPs are TM-polarized modes having an electric field perpendicular to the interface. The typical propagation constant β of the SPP supported at the interface of a semi-infinitely extended dielectric–metal interface is given by[56]

$$\beta = \frac{\omega}{c}\left(\frac{\varepsilon_1 \varepsilon_m}{\varepsilon_1 + \varepsilon_m}\right)^{\frac{1}{2}}, \tag{9.34}$$

where ω is the angular frequency of the incident wave, c is the speed of light, and ε_1 and ε_m are the dielectric constants of the dielectric and the metal, respectively.

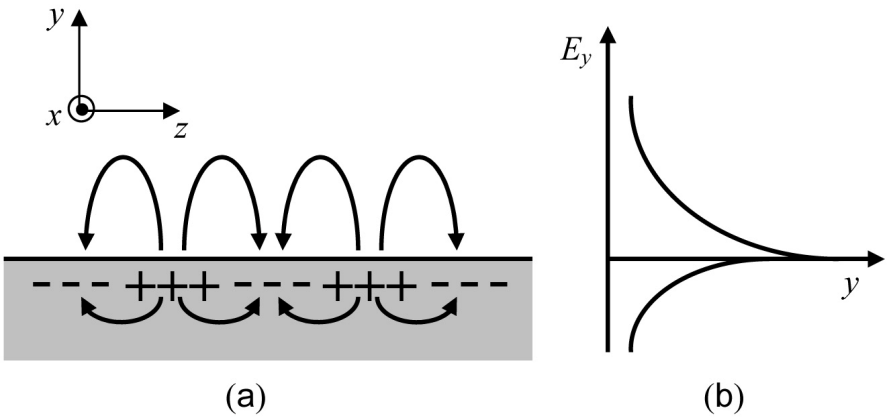


Figure 9.18 Schematic diagram of (a) the electron density wave propagating along a metal–dielectric interface, and (b) the exponentially decaying nature of its electromagnetic field with distance from the interface.

As mentioned earlier, SPPs have their field maxima near the interface; hence, they are extremely sensitive to any change in the characteristics of the medium in the vicinity of the interface. Using this property of SPPs, a large number of fiber optic refractive index sensors are reported in the literature.[57–59]

References

1. S. C. Rashleigh, "Origins and control of polarization effects in single-mode fibers," *J. Lightw. Technol.* **1**, 312–331 (1983).
2. D. N. Payne, A. J. Barlow, and J. J. R. Hansen, "Development of low- and high-birefringence optical fibers," *IEEE J. Quantum Electron.* **18**, 477–488 (1982).
3. C. Yeh, "Elliptical dielectric waveguides," *J. Appl. Phys.* **33**, 3235–3245 (1962).
4. R. B. Dyott, J. R Cozens, and D. G. Morris, "Preservation of polarization in optical fiber waveguides with elliptical cores," *Electron. Lett.* **15**, 380–382 (1979).
5. C. Yeh, K. Ha, S. B. Dong, and W. P. Brown, "Single-mode optical waveguides," *Appl. Opt.* **18**, 1490–1504 (1979).
6. K. Okamato, T. Hosaka, and Y. Sasaki, "Linearly single polarization fibers with zero polarization mode dispersion," *IEEE J. Quantum Electron.* **18**, 496–503 (1982).
7. J. D. Love, R. A. Sammut, and A. W. Synder, "Birefringence in elliptically deformed optical fibers," *Electron. Lett.* **15**, 615–616 (1979).
8. R. A. Sammut, "Birefringence in slightly elliptical optical fibers," *Electron. Lett.* **16**, 728–729 (1980).
9. A. Kumar, R. K. Varshney, and K. Thyagarajan, "Birefringence calculations in elliptical-core fibers," *Electron. Lett.* **20**, 112–113 (1984).
10. A. Kumar, K. Thyagarajan, and A. K. Ghatak, "Analysis of rectangular core dielectric waveguide: an accurate perturbation approach," *Opt. Lett.* **8**, 63–65 (1983).
11. S. R. Rengarajan and J. E. Lewis, "First higher-mode cutoff in two layer elliptical fiber waveguides," *Electron. Lett*. **16**, 263–264 (1980).
12. R. K. Varshney and A. Kumar, "Effect of depressed inner cladding on the characteristics of elliptical core fibers," *Opt. Lett.* **9**, 522–524 (1984).
13. W. Eickhoff, "Stress-induced single-polarization single-mode fiber," *Opt. Lett.* **7**, 629–631 (1982).
14. P. L. Chu, "Thermal stress induced birefringence in single-mode elliptical optical fiber," *Electron. Lett.* **18**, 45–47 (1982).
15. J. Sakai and T. Kimura, "Birefringence caused by thermal stress in elliptically deformed core optical fibers," *IEEE J. Quantum. Electron.* **18**, 1899–1909 (1982).

16. M. P. Varnham, D. N. Payne, A. J. Barlow, and R. D. Birch, "Analytic solution for the birefringence produced by thermal stress in polarization-maintaining optical fibers," *J. Lightw. Technol.* **1**, 332–339 (1983).
17. T. Okoshi and K. Oyamada, "Single polarization single mode optical fiber with refractive index pits on both sides of cores," *Electron. Lett.* **16**, 712–713 (1980).
18. T. Okoshi, K. Oyamada, M. Nishimura, and H. Yokota, "Side tunnel fiber: an approach to polarization maintaining optical waveguiding scheme," *Electron. Lett.* **18**, 824–826 (1982).
19. T. Okoshi, T. Aihara, and K. Kikuchi, "Prediction of the ultimate performance of side tunnel single polarization fiber," *Electron. Lett.* **19**, 1080–1082 (1983).
20. A. N. Kaul, A. Kumar, and K. Thyagarajan, "Polarization characteristics of side pit and side tunnel fibers using the effective index method," *IEEE J. Lightw. Technol.* **LT-5**, 1610–1612 (1987).
21. R. K. Varshney and A. Kumar, "Birefringence calculations in side tunnel optical fibers: a rectangular core waveguide modal," *Opt. Lett.* **11**, 45–47 (1986).
22. R. D. Birch, D. N. Payne, and M. P. Varnham, "Fabrication of polarization maintaining fiber using gas phase etching," *Electron Lett.* **18**, 1036–1038 (1982).
23. T. Hosaka, K. Okamoto, T. Miya, Y. Sasaki, and T. Edahiro, "Low loss single polarization fibers with asymmetrical strain birefringence," *Electron. Lett.* **17**, 530–531 (1981).
24. Y. Sasaki, K. Okamoto, T. Hosaka, and N. Shibata, "Polarization maintaining and absorption reducing fibers," paper ThCC6 in *Optical Fiber Communication 1982*, *OSA Technical Digest Series*, Optical Society of America (1982).
25. N. Shibata, Y. Sasaki, K. Okamoto, and T. Hosaka, "Fabrication of polarization maintaining and absorption reducing fibers," *J. Lightw. Technol.* **1**, 38–43 (1983).
26. K. Okamoto, T. Hosaka, and T. Edahiro, "Stress analysis of optical fibers by a finite element method," *IEEE J. Quantum Electron.* **17**, 2123–2129 (1981).
27. P. L. Chu and R. A. Sammut, "Analytical method for calculation of stresses and material birefringence in polarization maintaining optical fiber," *J. Lightw. Technol.* **2**, 650–662 (1984).
28. Y. Sasaki, T. Hosaka, M. Horiguchi, and J. Noda, "Design and fabrication of low-loss and low cross-talk polarization-maintaining optical fibers," *J. Lightw. Technol.* **LT-4**, 1097–1102 (1986).
29. A. Papp and H. Harns, "Polarization optics of index-gradient optical waveguide fibers," *Appl. Opt.* **14**, 2406–2411 (1975).

30. R. Ulrich and A. Simon, "Polarization optics of twisted single-mode fibers," *Appl. Opt.* **18**, 2241–2251 (1979).

31. A. N. Kaul, A. Kumar, and K. Thyagarajan, "Calculation of relative single polarization bandwidth of side pit fibers," *Electron. Lett.* **22**, 786–787 (1986).

32. T. Hosaka, "Single mode fiber with asymmetrical refractive index pits on both sides of core," *Electron. Lett.* **17**, 191–193 (1981).

33. A. W. Synder and F. Ruhl, "New single mode single polarization optical fiber," *Electron. Lett.* **19**, 185–186 (1983).

34. M. P. Varnham, D. N. Payne, R. D. Birch, and E. J. Tarbox, "Bend behavior of polarizing optical fibers," *Electron. Lett.* **19**, 679–680 (1983).

35. K. Okamoto, T. Hosaka, and J. Noda, "High birefringence polarizing fiber with flat cladding," *J. Lightw. Technol.* **LT-3**, 758–762 (1985).

36. J. Noda, K. Okamoto, and Y. Sasaki, "Polarization-maintaining fibers and their applications," *J. Lightw. Technol.* **LT-4**, 1071–1089 (1986).

37. J. Noda, I. Yokohama, and K. Okamoto, "Single-mode fiber devices," *Optoelectron. Dev. Technol.* **1**, 175–194 (1986).

38. J. N. Blake, S. Y. Huang, B. Y. Kim, and H. J. Shaw, "Strain effects on highly elliptic core fibers," *Opt. Lett.* **12**, 732–734 (1987).

39. K. A. Murphy, M. S. Miller, A. M. Vengsarkar, and R. O. Claus, "Elliptic-core two-mode optical fiber sensor implementation methods," *J. Lightw. Technol.* **8**, 1688–1696 (1990).

40. S. Y. Huang, J. N. Blake, and B. Y. Kim, "Perturbation effects on mode propagation in highly elliptic-core two-mode fibers," *J. Lightw. Technol.* **8**, 23–33 (1990).

41. F. A. Castro, S. R. M. Carneiro, O. Lisboa, and S. L. A. Carrara, "Two mode optical fiber accelerometer," *Opt. Lett.* **17**, 1474–1475 (1992).

42. B. Y. Kim, J. N. Blake, H. E. Ehgan, and H. J. Shaw, "All fiber acousto-optic frequency shifter," *Opt. Lett.* **11**, 389–391 (1986).

43. O. Lisboa and S. L. A. Carrara, "In line acousto-optic frequency shifter in two mode fiber," *Opt. Lett.* **17**, 154–156 (1992).

44. J. N. Blake, B. Y. Kim, and H. J. Shaw, "Fiber optic modal coupler using periodic microbending," *Opt. Lett.* **11**, 177–179 (1986).

45. C. D. Poole, K. T. Nelson, J. M. Wisenfield, and A. R. McCormic, "Broad-band dispersion compensation using the higher-order spatial-mode in a two-mode fiber," *Opt. Lett.* **17**, 985–987 (1992).

46. C. D. Poole, J. M. Wisenfield, and D. J. Digiovanni, "Elliptic-core dual-mode fiber dispersion compensator," *IEEE Photon. Technol. Lett.* **5**, 194–197 (1993).

47. A. Kumar, U. K. Das, and R. K. Varshney, "Design of a mode filter consisting of two dual mode highly elliptical core fibers," *J. Lightw. Technol.* **8**, 34–38 (1990).
48. A. Ghatak and K. Thyagarajan, *Optical Electronics*, Cambridge University Press, New York (1989).
49. K. Yokohama and J. Noda, "Fiber-optic polarizing beam splitter employing birefringent-fiber coupler," *Electron Lett.* **21**, 415–416 (1983).
50. A. Kumar and R. Ulrich, "Birefringence of optical fiber pressed into a V-groove," *Opt. Lett.* **6**, 644–646 (1981).
51. F. Maystre and A. Bertholds, "Zero-birefringence optical fiber holder,"*Opt. Lett.* **12**, 126–128 (1987).
52. H. C. Lefevre, "Single mode fiber fractional wave devices and polarization controllers," *Electron. Lett.* **16**, 778–780 (1980).
53. W. Johnstone, G. Stewart, T. Hart, and B. Culshaw, "Fiber optics polarizers and polarizing couplers," *Electron. Lett.* **24**, 868–870 (1988).
54. W. Johnstone, G. Stewart, T. Hart, and B. Culshaw, "Surface plasmon polaritons in thin metal films and their role in fiber optic polarizing devices," *J. Lightw. Technol.* **8**, 538–544 (1990).
55. R. H. Ritchie, "Plasma losses by fast electrons in thin films," *Phys. Rev.* **106**, 874–881 (1957).
56. H. Raether, *Surface Plasmons on Smooth and Rough Surfaces and on Gratings*, Springer Tracts in Modern Physics, vol. 111, Springer-Verlag, New York (1988).
57. J. Homola, *Surface Plasmon Resonance Based Sensors*, *Springer Series on Chemical Sensors and Biosensors*, vol. 4, Springer-Verlag, Berlin (2006).
58. G. Nemova and R. Kashyap, "Fiber Bragg grating assisted surface plasmon polariton sensor," *Opt. Lett.* **31**, 2118–2120 (2006).
59. S. M. Tripathi, A. Kumar, E. Marin, and J. P. Meunier, "Side-polished optical fiber grating-based refractive index sensors utilizing the pure surface plasmon polariton," *J. Lightw. Technol.* **26**, 1980–1985 (2008).

Chapter 10
Polarization Mode Dispersion in Optical Fibers

10.1 Introduction

In the previous chapter, we discussed devices based on deliberately introduced birefringence in an optical fiber, thus utilizing the polarization effects to our advantage. In Chapter 10, we discuss the detrimental aspect of birefringence in optical fibers—namely, polarization mode dispersion (PMD).

The existence of birefringence in a fiber implies that the fiber supports two orthogonally polarized modes that have different effective indices and hence propagate with different group velocities in the fiber. An optical pulse launched into such a fiber would be split into two orthogonally polarized pulses, which would then propagate with different propagation constants and group velocities. The two pulses thus reach the output end of the fiber at slightly different times and with different phases. The superimposition of these two pulses leads to the generation of an optical pulse that is now temporally more broadened as compared to the input pulse. Thus, the pulse becomes dispersed due to the effect of fiber birefringence, and the phenomenon is called polarization mode dispersion. PMD is a serious limitation in the case of ultrahigh-bit-rate (> 10 Gb) fiber communication links, as it puts a cap on the bit rate of the link as well as causes errors in data transmission. In the following, we discuss some basic concepts involved in the understanding of PMD in optical fibers.

10.2 PMD in Short-Length and High-Birefringence Fibers

In short-length SMFs or Hi-Bi fibers, one can assume that the birefringence is constant in magnitude as well as in direction so that there is no polarization mode coupling. In such cases, the two polarization modes—namely, slow and fast—are fixed, so PMD is completely deterministic and not random. If n_s and n_f are the effective indices of the slow and fast modes, respectively, the corresponding propagation constants will be given by

$$\beta_s = \frac{\omega}{c} n_s \quad \text{and} \quad \beta_f = \frac{\omega}{c} n_f , \tag{10.1}$$

where ω is the angular frequency. The two orthogonally polarized pulses will be traveling with different group velocities given by

$$\frac{1}{v_{gs}} = \frac{d\beta_s}{d\omega}, \quad \text{and} \quad \frac{1}{v_{gf}} = \frac{d\beta_f}{d\omega} . \tag{10.2}$$

Thus, the two pulses will take different times to travel a given length L of the fiber. The difference in the propagation times $\Delta\tau$ is known as the differential group delay (DGD) of the fiber and is given by

$$\Delta\tau = \frac{L}{v_{gs}} - \frac{L}{v_{gf}} = L\frac{d}{d\omega}[\beta_s - \beta_f] = L\Delta\beta' , \tag{10.3}$$

where $\Delta\beta = (\beta_s - \beta_f)$, and $\Delta\beta'$ represents the derivative of $\Delta\beta$ with respect to ω. The DGD is taken as the measure of PMD of a given fiber. Figure 10.1 shows the effect of the PMD on the input pulse in the time domain, demonstrating that an input pulse is divided into two orthogonally polarized pulses that are separated in time by $\Delta\tau$ at the output end.

It should be noted that in short-length fibers, the DGD grows linearly with fiber length [see Eq. (10.3)]. It is clear from Eq. (10.1) that the two polarized modes will accumulate a phase difference between them, which at a distance z from the input end is given by

$$\delta(\omega, z) = (\beta_s - \beta_f)z = \frac{\omega}{c}(n_s - n_f)z = \frac{\omega}{c}(\Delta n_{eff})z , \tag{10.4}$$

where $\Delta n_{eff} = (n_s - n_f)$. The phase difference $\delta(\omega, z)$ depends on both frequency and distance. As a result, the SOP of a polarized incident pulse will change with both distance and input frequency. Equation (10.3) can also be put in the form

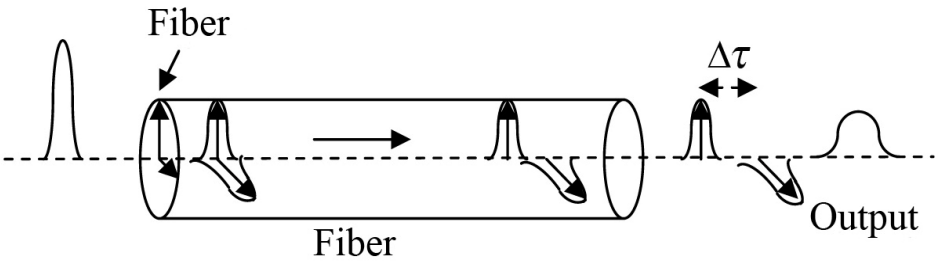


Figure 10.1 A schematic of pulse broadening due to PMD.

$$\Delta\tau = \frac{d}{d\omega}[L(\beta_s - \beta_f)] = \frac{d}{d\omega}\delta(\omega, L). \tag{10.5}$$

Equation (10.5) shows that $\Delta\tau$ is equal to the rate of change of $\delta(\omega, L)$ with respect to frequency ω. A change in the phase difference $\delta(\omega, L)$ means a change in the output SOP, so it is implied that $\Delta\tau$ can be obtained by knowing the rate of change of the output SOP with respect to the input frequency ω. For example, if the change $\Delta\omega$ in the input frequency ω is such that the output SOP repeats itself, the change in $\delta(\omega, L)$ will be equal to 2π, and hence, $\Delta\tau$ will be given by

$$\Delta\tau = \frac{2\pi}{\Delta\omega}. \tag{10.6}$$

This is the frequency domain picture of the PMD. In fact, as discussed later in this chapter, by measuring the rate of change of the output SOP with respect to the frequency (or wavelength) of input light, one can determine the PMD. This, in fact, is the working principle of several PMD measuring techniques.

It may be noted that Eq. (10.3) can also be recast as

$$\Delta\tau = L\frac{d}{d\omega}\left[\frac{\omega}{c}(n_s - n_f)\right] = \frac{L}{c}\left[\Delta n_{\text{eff}} + \omega\frac{d\Delta n_{\text{eff}}}{d\omega}\right]. \tag{10.7}$$

In a special case, if the dispersion of Δn_{eff} can be neglected, then

$$\Delta\tau = \frac{L}{c}\Delta n_{\text{eff}} = \frac{\lambda_0}{c}\frac{L}{L_b}, \tag{10.8}$$

where L_b represents the beat length of the fiber given by
$L_b = 2\pi / \Delta\beta = \lambda_0 / \Delta n_{\text{eff}}$.

Example 10.1 A birefringent fiber has 1-mm beat length at λ_0 = 1.5 μm. Assuming that the difference between the effective indices of the two polarized modes is independent of the wavelength, determine the DGD in a 1-m length of the fiber at (i) λ_0 = 1 μm and (ii) λ_0 = 1.5 μm.

Solution: The beat length ($\lambda_0 / \Delta n_{eff}$) at λ_0 = 1.5 μm is 1 mm. This means that $\Delta n_{\text{eff}} = \lambda_0 / L_b = 1.5\times10^{-3}$, and the DGD $\Delta\tau = \frac{L}{c}\Delta n_{\text{eff}} = \frac{1\times1.5\times10^{-3}}{3\times10^{8}} = 5\,\text{ps}$.
Since Δn_{eff} is independent of λ_0, the DGD will be same at λ_0 = 1 μm, i.e., 5 ps.

Example 10.2 Consider a piece of high-birefringence fiber in which a polarized light beam is coupled. It is observed that if the input wavelength is changed by 1 nm, the output SOP changes from a left-circular to a right-circular SOP. Obtain the DGD in the fiber at $\lambda_0 = 1$ μm.

Solution: In a highly birefringent fiber, the DGD is given by [see Eq. (10.5)]

$$\Delta\tau = \frac{\Delta\delta}{\Delta\omega},$$

where

$$\Delta\omega = -2\pi c\frac{\Delta\lambda_0}{\lambda_0^2}.$$

Thus,

$$\Delta\tau = \frac{\Delta\delta}{\Delta\omega} = -\frac{\lambda_0^2}{2\pi c}\frac{\Delta\delta}{\Delta\lambda_0}.$$

In the preceding three equations, $\Delta\omega$ is the change in the frequency corresponding to the wavelength change $\Delta\lambda_0$. The change in the output SOP from left circular to right circular means that $\Delta\delta = -\pi$. Substituting $\Delta\delta = -\pi$, $\Delta\lambda_0 = 10^{-9}$ m, and $\lambda_0 = 10^{-6}$ m, we obtain

$$\text{DGD} = \Delta\tau = \frac{10^{-12}}{2\pi\times 3\times 10^8}\times\frac{\pi}{10^{-9}} = 1.67\text{ps}.$$

10.3 PMD in Long-Length Fibers

In a long-distance fiber communication link, the fiber experiences stresses, bends, temperature changes, twists, etc., in a random fashion along the length of the link. Therefore, the birefringence along the fiber keeps changing both in magnitude as well as in direction. As a result, the birefringence is no longer additive; hence, the PMD does not grow linearly with the fiber length. Instead, it grows as a square root of the propagation distance, as will be discussed later in this section.

Due to random mode coupling, the calculation of PMD becomes very complicated. Fortunately, in a long-length fiber, there exist two orthogonal input polarization states, of which the output polarization states are independent of frequency to the first order. These input polarization states are known as the principal states of polarization (PSPs) of a given fiber length. Pulses launched in these polarization states emerge into two fixed output polarization states, which are also orthogonal to each other. Thus, if the input pulse is launched along one of the input PSPs, there is no splitting of the pulse, similar to the case with short-length or Hi-Bi fibers. The PSP model was proposed by Poole and Wagner[1] in 1986 and is extensively used in the calculation and understanding of PMD in long-length fibers.

It should be noted that due to random time variation of the birefringence along a long-length fiber link, PMD also varies randomly, so a statistical approach must be adopted when studying PMD. Using the coupled mode theory, Poole[2] has studied the effect of random mode coupling between the PSPs on the DGD and has arrived at an elegant expression for the root-mean-square (rms) DGD given by

$$\Delta\tau_{\text{rms}}(L) = \Delta\tau_0 \left(\sqrt{2}\,\frac{L_c}{L}\right)\left(e^{-L/L_c} - 1 + \frac{L}{L_c}\right)^{\frac{1}{2}}, \tag{10.9}$$

where $\Delta\tau_{\text{rms}} = \sqrt{\langle \Delta\tau^2 \rangle}$, L is the fiber length, and L_c is a constant known as the *coupling length* and is a measure of magnitude of the mode coupling along the fiber length. Further, $\Delta\tau_0$ is the DGD in the absence of mode coupling and hence will be given by Eq. (10.3), i.e., $\Delta\tau_0 = \Delta\beta' L$.

Equation (10.9) is valid for all values of L. For $L << L_c$,

$$\Delta\tau_{\text{rms}}(L) = \Delta\tau_0 \left(\sqrt{2}\,\frac{L_c}{L}\right)\left(\frac{L}{\sqrt{2}L_c}\right) = \Delta\tau_0 = \Delta\beta' L, \tag{10.10}$$

showing that the DGD varies linearly with L. On the other hand, for $L >> L_c$,

$$\Delta\tau_{\text{rms}}(L) = \Delta\tau_0 \left(\sqrt{2}\,\frac{L_c}{L}\right)\left(\frac{L}{L_c}\right)^{\frac{1}{2}} = \Delta\beta'\sqrt{2L_c L}, \tag{10.11}$$

where we have used $\Delta\tau_0 = \Delta\beta' L$. Equation (10.11) shows that in long-length fibers, the DGD grows as the square root of the length of the fiber. The parameter L_c is generally used to define the short-length and the long-length regimes of PMD as $L << L_c$ for the short-length regime, and $L >> L_c$ for the long-length regime.

Equation (10.11) suggests that PMD in a long-length fiber can be reduced by reducing $\Delta\beta'$ and L_c. Reducing $\Delta\beta'$ means that the birefringence in the fiber should be reduced, which obviously reduces PMD. Reducing L_c means that the coupling between the polarization modes should be increased. This is achieved in spun fibers, which are fabricated by spinning the fiber during the fiber-drawing process.[3,4] Such fibers were first fabricated for their applications in fiber sensors. Galtarossa and coworkers[5,6] have done extensive work on the PMD characteristics of spun fibers. Their studies have shown that periodic spinning in optical fibers is much more effective than uniform spinning for reducing the PMD in optical fibers.

10.4 Theoretical Modeling of PMD

As mentioned earlier, the birefringence along a long-length fiber keeps changing both in magnitude as well as in direction. As a result, the birefringence is no longer additive as in a constant-birefringence case such as a short-length fiber or a high-birefringence fiber. The PSP model, however, can be used to understand the PMD of a given long-length fiber in a similar fashion as the fixed polarization modes for a short-length or Hi-Bi fiber. Using the PSP model, Poole and coworkers[7] have derived a dynamical equation predicting the growth of the PMD along the fiber length. We will first discuss the concept of a PMD vector and the birefringence vector, which are used in the derivation of the dynamical equation discussed in Sec. 10.4.3.

10.4.1 PMD vector

Using the PSP model, in a given length z of the fiber, the DGD can be described by an equation similar to Eq. (10.5):

$$\Delta\tau = \frac{d}{d\omega}[z(\beta_{\mathrm{s}}^{\mathrm{psp}} - \beta_{\mathrm{f}}^{\mathrm{psp}})] = \frac{d}{d\omega}\delta^{\mathrm{psp}}(\omega, z), \tag{10.12}$$

where $\delta^{\mathrm{psp}}(\omega, z)$ represents the phase difference accumulated between the fast and slow PSPs of the fiber of length z, which is responsible for the change in the output SOP as ω changes. If the input frequency ω is varied, $\delta^{\mathrm{psp}}(\omega, z)$ will also vary; hence, on the Poincaré sphere, the output SOP will rotate on a circle whose rotation axis is parallel to the slow output PSP represented by $\hat{\mathbf{P}}_{\mathbf{psp}}^{-}$, as shown in Fig. 10.2. It is clear from Eq. (10.12) that the DGD is simply the rate of this rotation with ω. In view of this, the PMD of a given fiber length is generally described by a rotation vector in 3D Stokes subspace, defined as

$$\mathbf{\Omega} = \Delta\tau\,\hat{\mathbf{P}}_{\mathbf{psp}}^{-}, \tag{10.13}$$

where $\Delta\tau$ is the DGD, and $\mathbf{\Omega}$ is known as the PMD vector whose magnitude is the DGD and whose direction is parallel to the slow output PSP in the Stokes subspace. If the input frequency is changed by $\Delta\omega$, the change in $\delta^{\mathrm{psp}}(\omega, z)$ will be equal to $\Delta\tau\Delta\omega$. Thus, the change in the SOP at z can be obtained by rotating the sphere counterclockwise around $\mathbf{\Omega}$ by an angle $\Delta\tau\Delta\omega$ (see Fig. 10.2).

Let $\hat{\mathbf{S}}(\omega, z)$ and $\hat{\mathbf{S}}(\omega + \Delta\omega, z)$ represent the polarization states at distance z at frequencies ω and $\omega + \Delta\omega$, respectively. It is clear from Fig. 10.2, that the change $\Delta\hat{\mathbf{S}}$ in SOP will be such that

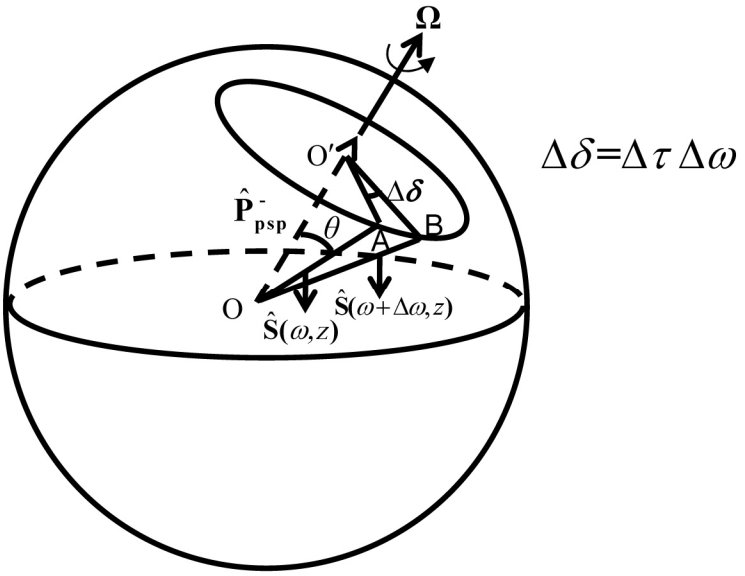


Figure 10.2 As frequency changes by $\Delta\omega$, the SOP at distance *z* along the fiber will rotate from point A to B on a circle by an angle $\Delta\delta$ with rotation axis parallel to $\mathbf{\Omega}$.

$$\left|\Delta\hat{\mathbf{S}}\right| = \left|\hat{\mathbf{S}}(\omega+\Delta\omega,z)-\hat{\mathbf{S}}(\omega,z)\right| = AB = O'A\Delta\delta = \Delta\tau\,\Delta\omega\sin\theta\,,$$

i.e.,

$$\Delta\hat{\mathbf{S}} = \Delta\omega\,\mathbf{\Omega}\times\hat{\mathbf{S}}(\omega,z)\,, \tag{10.14}$$

where θ is the angle between $\hat{\mathbf{P}}_{\mathbf{psp}}^{-}$ and $\hat{\mathbf{S}}(\omega,z)$.

Equation (10.14) shows that the evolution of the SOP along the fiber at *z* with frequency ω can be described by the following differential equation:

$$\frac{\partial\hat{\mathbf{S}}(\omega,z)}{\partial\omega} = \mathbf{\Omega}\times\hat{\mathbf{S}}(\omega,z)\,. \tag{10.15}$$

As one moves along the fiber length, the local birefringence changes the direction as well as the magnitude of the PMD vector. In order to include the effect of the local birefringence on the PMD vector in moving from z to $z+\Delta z$, we define another vector known as the *birefringence vector*.

10.4.2 Birefringence vector

As one moves from distance z to $z+\Delta z$, the SOP will be changed due to the local birefringence. At a given frequency ω, the phase difference accumulated between the two polarization modes of fiber length Δz will be given by

$$\Delta\delta'(z) = (\beta_s - \beta_f)\Delta z . \tag{10.16}$$

Here, β_s and β_f represent the propagation constants of the slow and the fast polarization modes of fiber section of length Δz. The change in the SOP can thus be obtained on the Poincaré sphere by rotating it counterclockwise by an angle $\Delta\delta'(z)$ around the Stokes vector, representing the local slow polarization mode (see Section 7.4). As the pulse moves along *z*, the rate of the previously mentioned rotation is given by

$$\frac{d\delta'(z)}{dz} = (\beta_s - \beta_f) = \Delta\beta . \tag{10.17}$$

Thus, at a given frequency ω, the evolution of the SOP of the pulse along the fiber length can be described by a rotation vector $\boldsymbol{\beta}$ (known as the birefringence vector), given by

$$\boldsymbol{\beta} = \Delta\beta\,\hat{\mathbf{P}}_{\mathbf{local}}^{-}, \tag{10.18}$$

where $\Delta\beta = (\beta_s - \beta_f)$, and $\hat{\mathbf{P}}_{\mathbf{local}}^{-}$ represents the Stokes vector of the local slow polarization mode of the fiber section of length Δz. Thus, on the Poincaré sphere, $\boldsymbol{\beta}$ represents a vector whose magnitude is $\Delta\beta$ and whose direction is along $\hat{\mathbf{P}}_{\mathbf{local}}^{-}$. Clearly, as the pulse moves along the fiber, its SOP will be moving on a circle with its rotation axis parallel to $\boldsymbol{\beta}$. The change in the SOP in moving a distance Δz along the fiber can thus be obtained from its SOP at *z* by rotating the sphere counterclockwise around $\boldsymbol{\beta}$ by an angle $\Delta\beta\Delta z$ (see Fig. 10.3).

Let $\hat{\mathbf{S}}(\omega, z)$ and $\hat{\mathbf{S}}(\omega, z+\Delta z)$ represent the polarization states at frequency ω at distances *z* and $z+\Delta z$, respectively. It is evident from Fig. 10.3 that the change $\Delta\hat{\mathbf{S}}$ in SOP will be such that

$$\left|\Delta\hat{\mathbf{S}}\right| = \left|\hat{\mathbf{S}}(\omega, z+\Delta z) - \hat{\mathbf{S}}(\omega, z)\right| = A'B' = \Delta\beta\,\Delta z \sin\theta' ,$$

i.e.,

$$\Delta\hat{\mathbf{S}} = \Delta z\,\boldsymbol{\beta} \times \hat{\mathbf{S}}(\omega, z) . \tag{10.19}$$

Thus, the evolution of the SOP along the fiber length at frequency ω can be described by the following differential equation:

$$\frac{\partial\hat{\mathbf{S}}(\omega, z)}{\partial z} = \boldsymbol{\beta} \times \hat{\mathbf{S}}(\omega, z) . \tag{10.20}$$

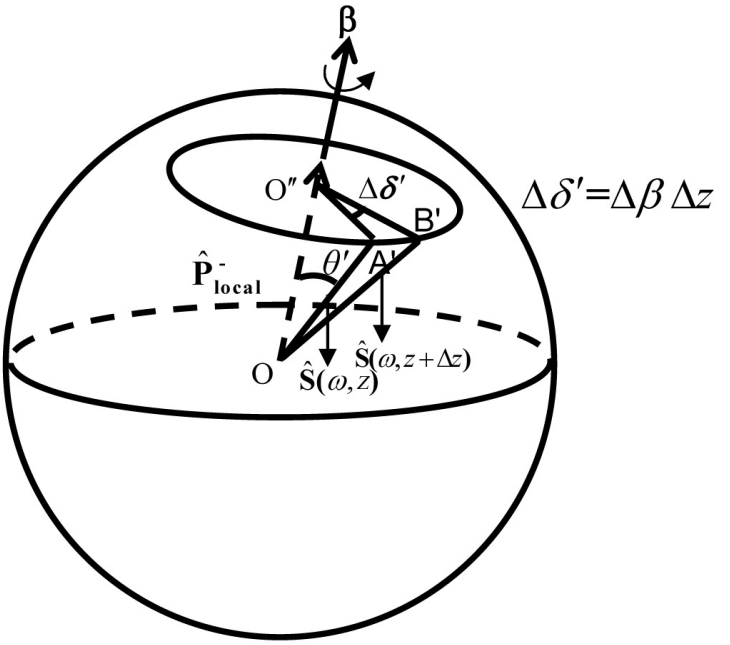


Figure 10.3 As distance changes by Δz, the SOP will rotate from point A' to B' on a circle by an angle $\Delta\delta'$ with rotation axis parallel to $\boldsymbol{\beta}$.

It may be mentioned here that in the case of long-length fibers ($L >> L_c$), the directions of $\boldsymbol{\Omega}$ and $\boldsymbol{\beta}$ are generally different, both varying with *z*. On the other hand, for a short-length fiber, since the birefringence is constant, and there is no polarization mode coupling along the fiber length, the directions of the $\boldsymbol{\Omega}$ and $\boldsymbol{\beta}$ vectors will be parallel to each other.

10.4.3 Dynamical equation for the PMD vector

A dynamical equation[7] predicting the growth of the PMD vector along the fiber length can be derived using Eqs. (10.15) and (10.20). Assuming the continuity of the PMD vector, one can write

$$\frac{\partial}{\partial z}\left(\frac{\partial \hat{\mathbf{S}}(\omega, z)}{\partial \omega}\right) = \frac{\partial}{\partial \omega}\left(\frac{\partial \hat{\mathbf{S}}(\omega, z)}{\partial z}\right). \tag{10.21}$$

Substituting $\partial\hat{\mathbf{S}}/\partial\omega$ from Eq. (10.15) and $\partial\hat{\mathbf{S}}/\partial z$ from Eq. (10.20) into the preceding equation, one obtains

$$\frac{\partial \boldsymbol{\Omega}(\omega, z)}{\partial z} \times \hat{\mathbf{S}}(\omega, z) + \boldsymbol{\Omega}(\omega, z) \times \frac{\partial \hat{\mathbf{S}}(\omega, z)}{\partial z} = \frac{\partial \boldsymbol{\beta}(\omega, z)}{\partial \omega} \times \hat{\mathbf{S}}(\omega, z) + \boldsymbol{\beta}(\omega, z) \times \frac{\partial \hat{\mathbf{S}}(\omega, z)}{\partial \omega}.$$

Substituting again $\partial\hat{\mathbf{S}}/\partial\omega$ and $\partial\hat{\mathbf{S}}/\partial z$ in the preceding equation, and using the vector identity $\mathbf{A}\times(\mathbf{B}\times\mathbf{C}) = \mathbf{B}(\mathbf{A}\cdot\mathbf{C}) - \mathbf{C}(\mathbf{A}\cdot\mathbf{B})$, one obtains

$$\frac{\partial\boldsymbol{\Omega}(\omega,z)}{\partial z}\times\hat{\mathbf{S}}(\omega,z)=\frac{\partial\boldsymbol{\beta}(\omega,z)}{\partial\omega}\times\hat{\mathbf{S}}(\omega,z)+\left[\boldsymbol{\beta}(\omega,z)\times\boldsymbol{\Omega}(\omega,z)\right]\times\hat{\mathbf{S}}(\omega,z),$$

i.e.,

$$\frac{\partial\boldsymbol{\Omega}(\omega,z)}{\partial z} = \frac{\partial\boldsymbol{\beta}(\omega,z)}{\partial\omega} + \boldsymbol{\beta}(\omega,z)\times\boldsymbol{\Omega}(\omega,z). \tag{10.22}$$

Equation (10.22) is the required dynamical equation that describes the growth of the PMD vector along the fiber length. It must be noted that the first term on the right-hand side is the local DGD per unit length, which is the deriving term responsible for the growth of PMD. The second term accounts for the change of the PMD vector due to local birefringence, because of which the DGD does not grow linearly with the fiber length.

Poole and coworkers[7] have solved the preceding equation using Monte Carlo simulations, taking the random variation of birefringence into account. Their results show that the probability density function for the magnitude of the PMD in a long-length fiber follows a Maxwellian distribution,[7,8] given by

$$P(\Delta\tau)=\frac{32}{\pi^2}\frac{\Delta\tau^2}{\langle\Delta\tau\rangle^3}\exp\left(-\frac{4\Delta\tau^2}{\pi\langle\Delta\tau\rangle^2}\right), \tag{10.23}$$

where $\langle\Delta\tau\rangle$ is the mean DGD.

10.4.4 Concatenation model: an alternative approach

An alternative approach to understanding the PMD characteristics of a long-length transmission fiber of length L is to consider the given fiber as a concatenation of an N number of small, linearly birefringent fiber segments whose lengths are randomly chosen around a certain mean length. Each segment has a fixed set of birefringent axes; however, various sections are oriented randomly with respect to the other sections.[9,10] Figure 10.4 shows the schematic of the concatenation model.

The Jones matrix T_n for the n'th section can be written as

$$T_n=\begin{pmatrix}\cos\theta_n & -\sin\theta_n\\ \sin\theta_n & \cos\theta_n\end{pmatrix}\begin{pmatrix}\exp(i\frac{Bh_n}{2}) & 0\\ 0 & \exp(-i\frac{Bh_n}{2})\end{pmatrix}\begin{pmatrix}\cos\theta_n & \sin\theta_n\\ -\sin\theta_n & \cos\theta_n\end{pmatrix}, \tag{10.24}$$

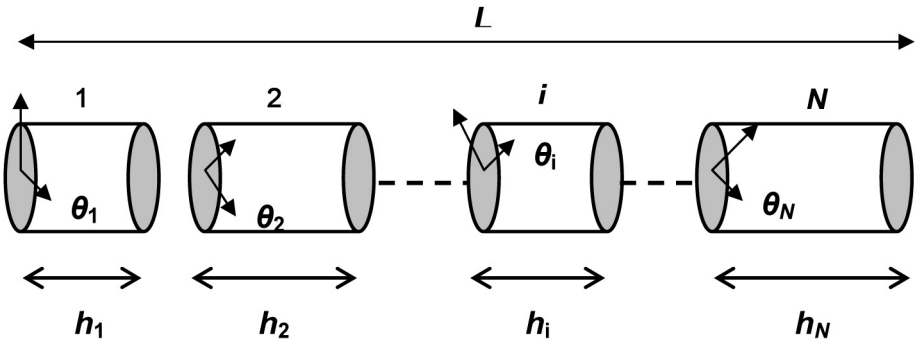


Figure 10.4 Schematic showing the fiber as a concatenation of numerous small birefringent sections of random lengths and orientations.

where $B = (\beta_s - \beta_f)_n$ represents the birefringence, θ_n is the orientation of the fast axis of the n'th segment with respect to a fixed laboratory axis (for instance, the x axis), and h_n is the length of the n'th segment. Thus, the Jones matrix corresponding to the entire length of fiber can be expressed as a multiplication over T_n for all of the N sections:

$$T = \prod_{n=1}^{N} T_n \,. \tag{10.25}$$

To calculate the DGD, the method of Jones matrix eigenanalysis is used, which is discussed in the following section.

10.4.4.1 Jones matrix eigenanalysis

In this technique, we first obtain the Jones matrix $T(\omega_1)$ and $T(\omega_2)$ of the given fiber length L, predicting the output polarization states corresponding to a given input SOP $\hat{\mathbf{S}}_\mathbf{i}$ at two closely spaced frequencies ω_1 and ω_2. If $\hat{\mathbf{S}}_\mathbf{o}(\omega_1, L)$ and $\hat{\mathbf{S}}_\mathbf{o}(\omega_2, L)$ are the corresponding output polarization states,

$$\hat{\mathbf{S}}_\mathbf{o}(\omega_1, L) = T(\omega_1)\hat{\mathbf{S}}_\mathbf{i}\,, \tag{10.26}$$

and

$$\hat{\mathbf{S}}_\mathbf{o}(\omega_2, L) = T(\omega_2)\hat{\mathbf{S}}_\mathbf{i}\,. \tag{10.27}$$

Using Eqs. (10.26) and (10.27), we then obtain the matrix relating $\hat{\mathbf{S}}_\mathbf{o}(\omega_1, L)$ to $\hat{\mathbf{S}}_\mathbf{o}(\omega_2, L)$ as

$$\hat{\mathbf{S}}_\mathbf{o}(\omega_2, L) = \Gamma(\bar{\omega}, \Delta\omega)\, \hat{\mathbf{S}}_\mathbf{o}(\omega_1, L)\,, \tag{10.28}$$

where $\bar{\omega} = (\omega_1 + \omega_2)/2$, $\Delta\omega = (\omega_2 - \omega_1)$, and $\Gamma(\bar{\omega}, \Delta\omega) = T(\omega_2)T^{-1}(\omega_1)$ describes the evolution of the output SOP with frequency. The eigenstates of $\Gamma(\bar{\omega}, \Delta\omega)$ are the two PSPs, and hence, its eigenvalues are $\rho_s = \exp(-i\tau_s \Delta\omega)$ and $\rho_f = \exp(-i\tau_f \Delta\omega)$, respectively, corresponding to the slow and fast eigenvectors. The DGD is then given by[11]

$$\Delta\tau(\bar{\omega}) = \tau_s - \tau_f = \left| \frac{\mathrm{Arg}(\rho_s / \rho_f)}{\Delta\omega} \right|, \tag{10.29}$$

where τ_s and τ_f are the group delay times of the pulses traveling along the slow and fast PSPs of the fiber, and $\mathrm{Arg}(\,)$ is the argument function. Here, the value of $\Delta\omega$ is chosen such that the product $\Delta\tau\Delta\omega$ is sufficiently small to avoid any omission of full-2π rotations in the $\mathrm{Arg}(\,)$ function.

To obtain the probability distribution of the DGD, $\Delta\tau$ is calculated a large number of times by selecting θ_n to be a random quantity uniformly distributed in [0, π] and h_n to be a random quantity chosen from a Gaussian distribution around the mean length $h_m = L/N$ with a standard deviation Δh. These values of $\Delta\tau$ are arranged in the form of a relative frequency distribution, which is then fitted to a Maxwellian distribution of the form

$$P(\Delta\tau) = \frac{32}{\pi^2} \frac{\Delta\tau^2}{\langle \Delta\tau \rangle^3} \exp\left(-\frac{4\Delta\tau^2}{\pi \langle \Delta\tau \rangle^2} \right), \tag{10.30}$$

where $\langle \Delta\tau \rangle$ is the mean DGD. The mean PMD is then calculated as

$$D_{\mathrm{PMD}} = \frac{\langle \Delta\tau \rangle}{\sqrt{L}}. \tag{10.31}$$

10.5 PMD Measuring Techniques

The PMD of a given device is completely known if we know its polarization vector $\mathbf{\Omega}(\omega)$ as a function of frequency ω. As discussed earlier, the magnitude of $\mathbf{\Omega}(\omega)$ is equal to the DGD of the device, and its direction is the direction of the slow output PSP represented by the vector $\hat{\mathbf{P}}_{\mathbf{psp}}$ in the 3D Stokes subspace. Thus, the measurement of PMD is basically the measurement of the PMD vector $\mathbf{\Omega}(\omega)$. The measurement techniques can broadly be classified as (i) time-domain techniques and (ii) frequency-domain techniques.

A particular technique can be considered to be a time-domain technique provided $\tau_c << \Delta\tau$ and as a frequency-domain technique if $\tau_c >> \Delta\tau$, where τ_c

represents the coherence time of the light used in the measurement and $\Delta\tau$ is the DGD of the device under test (DUT). Various PMD measurement techniques are discussed in a recent book by Galtarossa and Menyuk.[12] In the following two sections, we discuss the working principle of these techniques briefly.

10.5.1 Time-domain techniques

The most simple and intuitive time-domain technique involves measuring τ_s and τ_f directly and calculating DGD as $\tau_s - \tau_f$. The corresponding measurement apparatus is also simple in principle and is shown schematically in Fig. 10.5. Short, polarized pulses are launched into the DUT, and the times of arrival of the two output pulses traveling along the slow and fast PSPs are measured using a fast oscilloscope. The input SOP is changed continuously using a polarization controller, and when it matches either of the input PSPs, only one output pulse will emerge. By measuring the polarization state of the slow output pulse, one can also obtain the $\hat{\mathbf{P}}_{\mathbf{psp}}^{-}$. One important requirement of this technique is that the width of the input pulses be extremely small in order to achieve the desired DGD resolution.

10.5.2 Frequency-domain techniques

The working principle of various frequency-domain techniques uses the fact that as the input frequency is changed, the change in the output SOP is described by

$$\frac{\partial \hat{\mathbf{S}}(L,\omega)}{\partial \omega} = \mathbf{\Omega} \times \hat{\mathbf{S}}(L,\omega)\,. \tag{10.32}$$

As discussed earlier, the preceding equation means that as the input frequency changes, the output SOP $\hat{\mathbf{S}}(L,\omega)$ rotates on the Poincaré sphere, around the PMD vector $\mathbf{\Omega}$. The rate of this rotation is given by

$$\left|\frac{d\delta}{d\omega}\right| = |\Omega| = \Delta\tau \;=\; \text{DGD}\,. \tag{10.33}$$

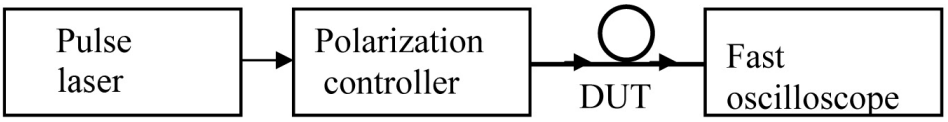


Figure 10.5 Schematic experimental setup used to measure PMD using time-domain techniques.

In Eq. (10.33), $d\delta$ represents the change in the phase difference accumulated between the output PSPs of the DUT due to a change in frequency $d\omega$. Thus, all frequency-domain techniques measure the variation in the output SOP with respect to the input frequency. The various frequency-domain techniques are thus basically polarimetric techniques. The schematic experimental setup in all of such techniques is similar and is shown in Fig. 10.6. Continuous and polarized light from a tunable laser with an appropriate polarization state (so that light is coupled in both of the PSPs) controlled by a polarization controller is passed through the DUT. At the output end, the SOP is measured as a function of input frequency using a polarimeter.

The experimental data obtained using frequency-domain techniques can be used to obtain the PMD vector $\mathbf{\Omega}$ by one of the following methods:

(i) Jones matrix eigenanalysis (JME),
(ii) Mueller matrix method (MMM), or
(iii) Poincaré sphere analysis (PSA).

Accordingly, frequency-domain techniques are also classified as JME, MMM, and PSA methods. The JME method was briefly discussed in Sec. 10.4.4.1; the working principles of the other two methods are similar.[12]

There are other methods that do not use JME, MMM, or PSA methods. In the next section, we discuss a polarimetric technique that is extremely simple to implement. This scheme measures only the DGD and is known as the wavelength-scanning method or the fixed-analyzer method.[13]

10.5.3 Wavelength-scanning method

The output SOP of a device with fixed eigen SOPs, such as a length of a high-birefringence fiber, traces a circle on the Poincaré sphere as the wavelength is changed. The axis of rotation will be parallel to the slow polarization mode $\hat{\mathbf{P}}_s$ of the fiber [Fig. 10.7(a)]. However, in a more general case, the birefringence may be due to a variety of factors—stress, twists, bends, etc.,—whose direction and magnitude may change randomly along the fiber length and with time. In such a case, the output SOP wanders on the sphere in a random fashion, as shown in Fig. 10.7(b).

The changes in the output polarization state are transformed into variations in the output light intensity by putting a "fixed analyzer" at the output, followed by an optical power meter. A schematic setup for the wavelength-scanning method for the measurement of PMD is shown in Fig. 10.8.

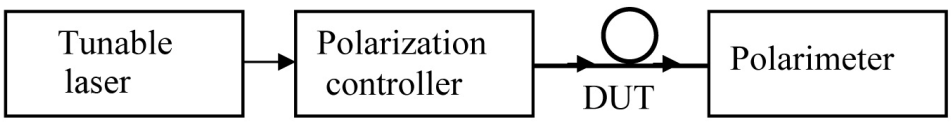


Figure 10.6 Schematic experimental setup used to measure PMD using frequency-domain techniques.

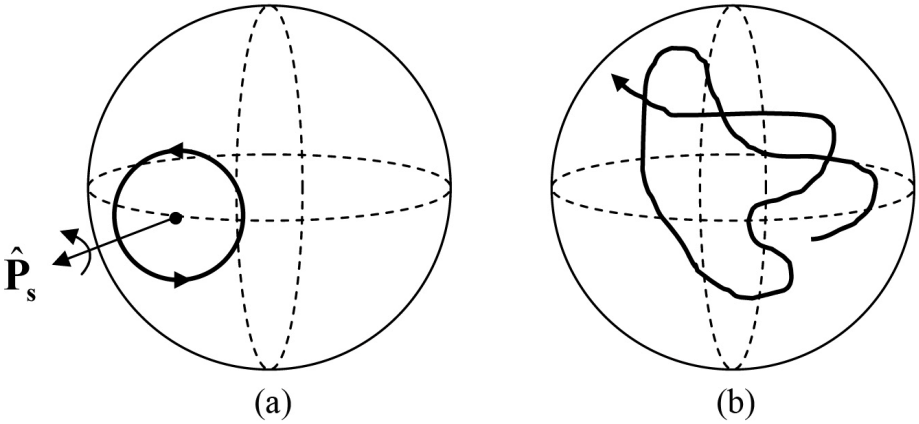


Figure 10.7 Evolution of the output SOP as the input wavelength is scanned in the case of (a) Hi-Bi and (b) practical single-mode fibers.

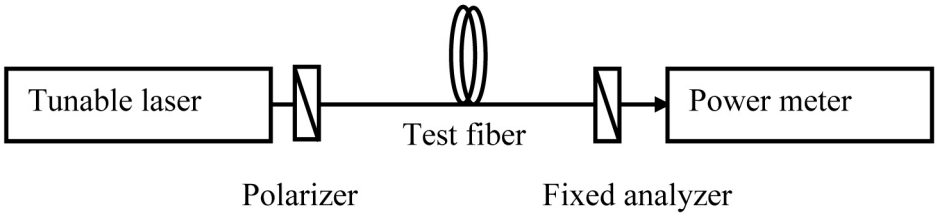


Figure 10.8 Schematic experimental setup used in the wavelength-scanning method.

It is easy to understand that in the case of a Hi-Bi fiber, the output intensity will vary periodically (as shown in Fig. 10.9) with wavelength, and each period will correspond to the change in the phase difference between the two polarization modes as $\Delta\delta = 2\pi$. Thus, by counting the number of times the intensity crosses the mean level, one can determine the DGD. For example, if the number of full cycles in the intensity variation between wavelengths λ_1 and λ_2 is *N*, then

$$\Delta\delta = 2N\pi \text{, and } \Delta\omega = c\,2\pi\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right). \tag{10.34}$$

Using Eq. (10.5),

$$\Delta\tau = \frac{\Delta\delta}{\Delta\omega} = \frac{N\lambda_1\lambda_2}{c\left(\lambda_2 - \lambda_1\right)}. \tag{10.35}$$

In the case of a Hi-Bi or a short-length fiber, there is no polarization mode coupling, and the output intensity variation is periodic with a fixed amplitude and period, as discussed earlier. However, in long-length single-mode fibers, due to random variation in the birefringence and finite coupling between orthogonal polarization states, the output intensity variation looks somewhat like that shown in Fig. 10.10.

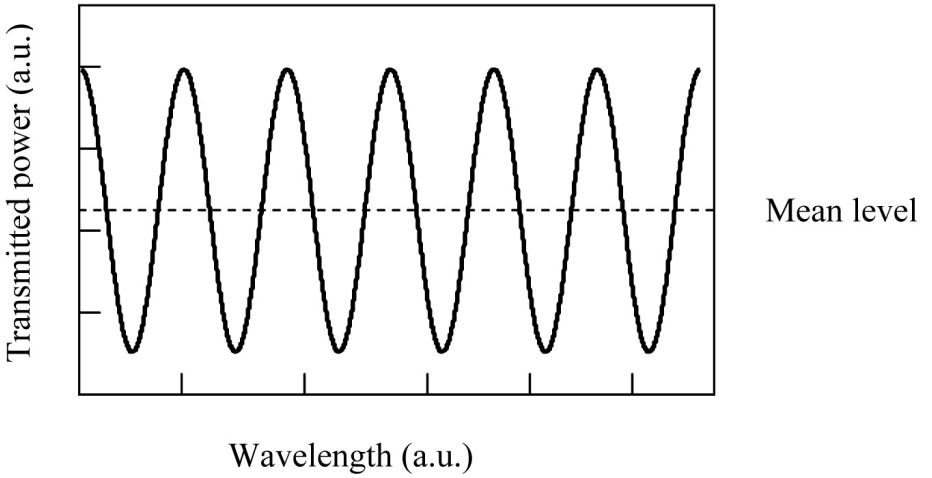


Figure10.9 Schematic representation of the variation in output intensity with respect to wavelength in a wavelength-scanning technique for a Hi-Bi or short-length fiber.

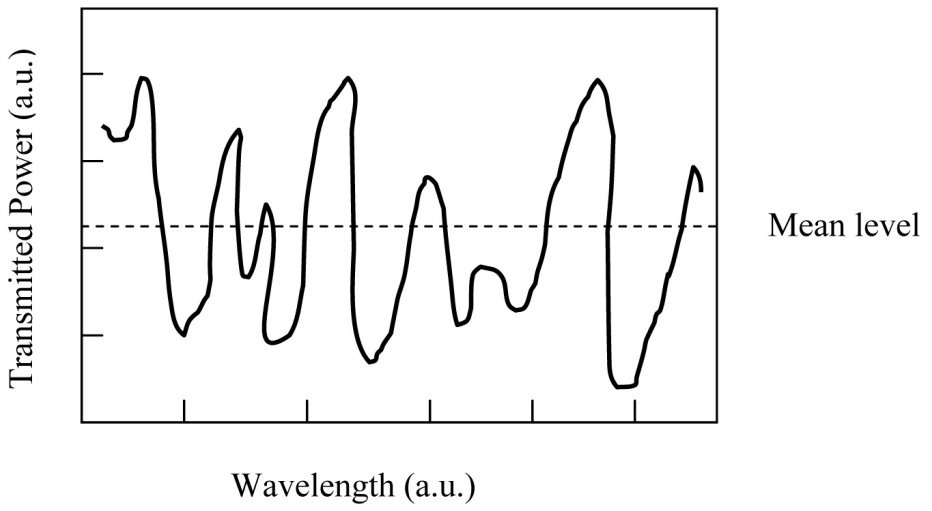


Figure 10.10 Schematic representation of the variation in output intensity with respect to wavelength in a wavelength-scanning technique for a long-length single-mode fiber.

By counting the number of oscillations in intensity about a mean threshold level, the value of DGD can still be estimated and is given by[13,14]

$$\Delta\tau = \frac{N\,k\,\lambda_1\lambda_2}{c\left(\lambda_2 - \lambda_1\right)}, \tag{10.36}$$

where constant k depends on the mode coupling, and $k = 0.81$ is recommended for highly mode-coupled fibers.[14] For standard communication-grade fibers, the PMD is typically 0.01 to 0.1 ps/km$^{1/2}$. It may be mentioned that, due to mode coupling, PMD in long lengths of fibers is proportional to the square root of the fiber length.

10.6 PMD Mitigation Techniques

Polarization-induced impairments are the main factor limiting the bit-rate per channel in a long-haul communication system. Furthermore, because of its statistical nature, one cannot fix an upper limit on the penalty imposed due to the PMD. Whatever limit one fixes, there is always a finite amount of probability (known as the *outage probability*) that the PMD-induced penalty will exceed the limit. The maximum tolerable limit of PMD is considered to be 10 to 20% of the bit duration. In an ultra-long-haul communication system, this limit is readily reached, due to the presence of chromatic dispersion compensators and optical amplifier modules. Thus, mitigating the PMD-induced impairments is extremely important. In this regard, research efforts are focused on two different goals: (i) producing low-PMD fibers and (ii) developing PMD compensators. In the following sections, we briefly discuss these two aspects of PMD mitigation.

10.6.1 Low-PMD fibers

As discussed earlier, an ideal single-mode fiber does not have any birefringence and hence, no PMD. Real fibers, however, have some core ellipticity as well as asymmetrical stresses present in the core that make the fiber birefringent, both of which are sources of PMD. Since it is extremely difficult to completely eliminate this birefringence, one has to find ways of reducing its contribution toward PMD. Research efforts made in this direction during the last two decades have established that spinning the fiber at very high speed while it is in the molten stage can significantly reduce the PMD.[3,4] It should be understood that spin is different from twist. Twist is introduced in the cold fiber, producing torsional stress and introducing a circular birefringence in the fiber through the photoelastic effect.[15,16] On the other hand, spinning the fiber in the molten state introduces only the rotation of the birefringence axes of the fiber, without introducing any torsional stress or circular birefringence in the fiber.

The working principle of PMD reduction through spinning can be understood by considering the fiber as various birefringent sections of equal length but rotated with respect to each other. Consider the case in which each section is

rotated by an angle of π/2 with respect to the previous section. Now, the fast (slow) axis of each section will be parallel to the slow (fast) axis of the following section. As a result, the PMD accumulated in one section will be exactly compensated for in the next section.

Both unidirectional as well as periodic spin functions are reported to obtain low-PMD fibers.[17–19] In order to obtain the exact factor by which the PMD is reduced due to spin, one has to solve the dynamical equation (10.21). In the case of unidirectional spin, Schuh et al.[17] have reported the analytical solution for the dynamical equation and have shown that due to the spin, the PMD is reduced by a spin-induced reduction factor (SIRF), given by

$$\text{SIRF} \simeq \frac{1}{\sqrt{1+4\left(\frac{L_\text{b}}{p}\right)^2}}, \tag{10.37}$$

where L_b represents the beat length of the unspun fiber, and p is the spin period, i.e., the fiber length over which the fiber completes one rotation. To gain a sense of the numerical value of the SIRF, if we consider L_b = 10 m, and p = 1 m, then SIRF ≈ 0.05.

10.6.2 PMD compensators

The second approach to PMD mitigation is to use a PMD compensator. It should be mentioned that it is extremely difficult to completely compensate for PMD, as it is a complex phenomenon due to its randomly varying nature. In order to understand the basic concepts involved in PMD-compensation techniques, we first discuss what is known as first- and second-order PMD.

10.6.2.1 First- and second-order PMD

PMD of a fiber can be described by the PMD vector $\boldsymbol{\Omega}$, which is given by $\boldsymbol{\Omega} = \Delta\tau\,\hat{\mathbf{P}}_{\mathbf{psp}}$ [see Eq. (10.13)], and which is frequency dependent. If we perform a Taylor series expansion of $\boldsymbol{\Omega}(\omega)$ around the carrier frequency ω_0, we obtain

$$\boldsymbol{\Omega}(\omega_0+\Delta\omega) = \boldsymbol{\Omega}(\omega_0)+\boldsymbol{\Omega}'(\omega_0)\Delta\omega+\ldots, \tag{10.38}$$

where

$$\boldsymbol{\Omega}'(\omega_0)=\left.\frac{d\boldsymbol{\Omega}(\omega)}{d\omega}\right|_{\omega_0}.$$

In the preceding equation, the first and second terms on the right-hand side are known as the first- and second-order PMD, respectively.[20] There will be higher-

order terms as well, which are known as higher-order PMDs. It should be noted that the first-order PMD is frequency independent.

The second-order PMD is proportional to $\boldsymbol{\Omega'(\omega_0)}$ and can be written as

$$\left.\frac{d\boldsymbol{\Omega}(\omega)}{d\omega}\right|_{\omega_0} = \left.\frac{d(\Delta\tau)}{d\omega}\right|_{\omega_0} \hat{\mathbf{P}}_{\mathbf{psp}}^{-}(\omega_0) + \Delta\tau(\omega_0)\left.\frac{d\,\hat{\mathbf{P}}_{\mathbf{psp}}^{-}}{d\omega}\right|_{\omega_0}. \qquad (10.39)$$

Thus, the second-order PMD consists of two terms: the first term is parallel to $\hat{\mathbf{P}}_{\mathbf{psp}}^{-}(\omega_0)$, i.e., $\boldsymbol{\Omega(\omega_0)}$, and the second term is perpendicular to $\boldsymbol{\Omega(\omega_0)}$, since $d\,\hat{\mathbf{P}}_{\mathbf{psp}}^{-}/d\omega$ is perpendicular to $\hat{\mathbf{P}}_{\mathbf{psp}}^{-}$. Neglecting higher orders of PMD, Fig. 10.11 schematically represents the PMD vector $\boldsymbol{\Omega(\omega)}$ and its first- and second-order PMD components, where vectors $\overrightarrow{\mathbf{AB}}$, $\overrightarrow{\mathbf{BC}}$ represent the first- and second-order PMD components of the total PMD vector $\overrightarrow{\mathbf{AC}}$, while vectors $\overrightarrow{\mathbf{BD}}$ and $\overrightarrow{\mathbf{DC}}$, respectively, represent the component of the second-order PMD that is parallel and perpendicular to the first-order PMD.

Various schemes are reported in the literature for compensating the first- and second-order PMD components. First-order PMD compensators are relatively easy to realize, the basic working principle of which is discussed briefly in the next section.

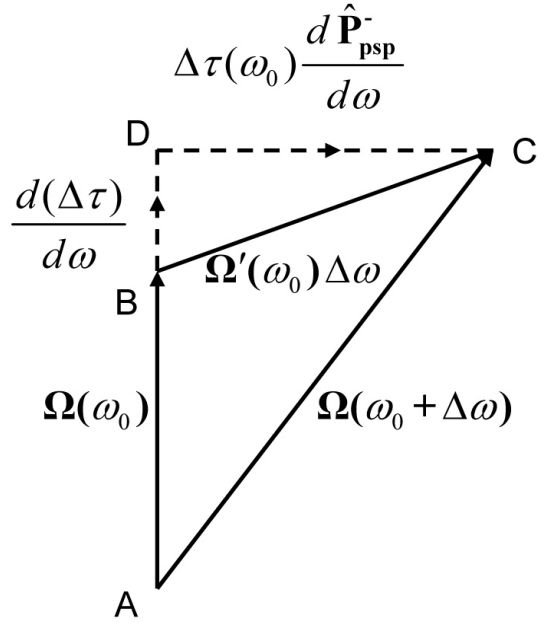


Figure 10.11 Schematic representation of the PMD vector $\boldsymbol{\Omega(\omega)}$ and its first- and second-order PMD components.

10.6.2.2 Working principle of a first-order PMD compensator

It should be clear from the preceding discussion that to the first-order approximation, PMD splits the input pulse into two pulses polarized along fast and slow PSPs. At the output end, these pulses are separated by a time delay that is equal to the DGD of the fiber. Both the directions of the PSPs as well as the magnitude of the DGD keep changing with time.

It must also be understood that for a given value of DGD, the PMD-induced impairments will be maximized when both PSPs carry equal power. For example, if only one PSP is excited, there will be no PMD-induced signal distortion.[21] Similarly, if the DGD is equal to zero, i.e., both pulses reach the output end at the same time, again there will be no PMD-induced signal distortion. On the Poincaré sphere, this means that if the input polarization state $\hat{\mathbf{S}}_\mathbf{i}$ is parallel to either $\boldsymbol{\Omega}(\omega)$ or $|\boldsymbol{\Omega}(\omega)|=0$, there will be no PMD-induced impairments, which is what a first-order PMD compensator is designed to achieve. Accordingly, there are two points of operation of a first-order PMD compensator, as discussed in the remainder of the section.

A PMD compensator[22] consists of a Hi-Bi fiber, a polarization controller (PC); a detector measuring the quality of the signal, (e.g., DOP), and a control algorithm that sends a feedback signal to the PC, as shown in Fig. 10.12. The compensator module is inserted at the receiver end of the fiber link. If the PMD of the fiber link and that of the compensator are denoted by $\boldsymbol{\Omega}_\mathbf{L}$ and $\boldsymbol{\Omega}_\mathbf{C}$, respectively, the overall PMD of the system (including the compensator) $\boldsymbol{\Omega}_\mathbf{Tot}$ will be given by $\boldsymbol{\Omega}_\mathbf{Tot} = \boldsymbol{\Omega}_\mathbf{L} + \boldsymbol{\Omega}_\mathbf{C}$ (see Fig. 10.13).

After inserting the compensator in the link, the PMD-induced impairments will be decided by the overall DGD = $|\boldsymbol{\Omega}_\mathbf{Tot}|$ and the angle θ_{Tot} between $\hat{\mathbf{S}}_\mathbf{i}$ (input SOP) and the $\boldsymbol{\Omega}_\mathbf{Tot}$ (overall PMD) shown in Fig. 10.13. By making either the overall DGD or the angle θ_{Tot} equal to zero, the PMD of the fiber link can be completely compensated. These two points of operation of a PMD compensator are shown in Fig. 10.14 and defined here:

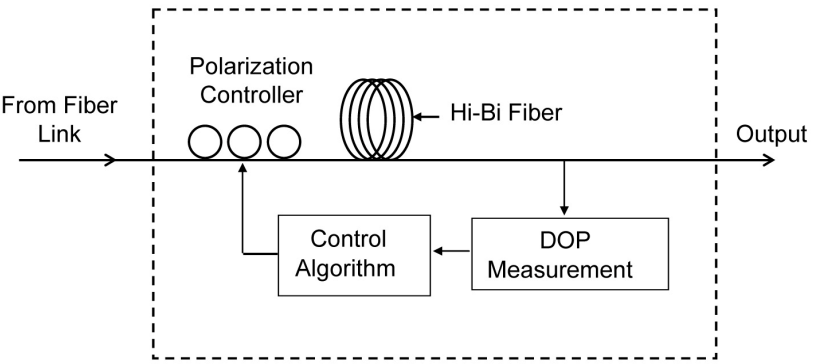


Figure 10.12 Schematic of a first-order PMD compensator.

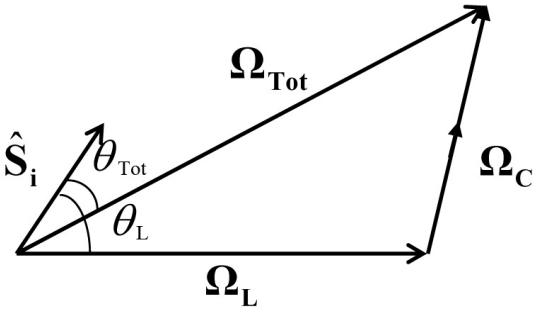


Fig. 10.13 Schematic vector representation of the PMD of the fiber link $\boldsymbol{\Omega}_{\mathbf{L}}$, the compensator $\boldsymbol{\Omega}_{\mathbf{C}}$, and the total PMD $\boldsymbol{\Omega}_{\mathbf{Tot}}$ with respect to the input polarization state $\hat{\mathbf{S}}_{\mathbf{i}}$.

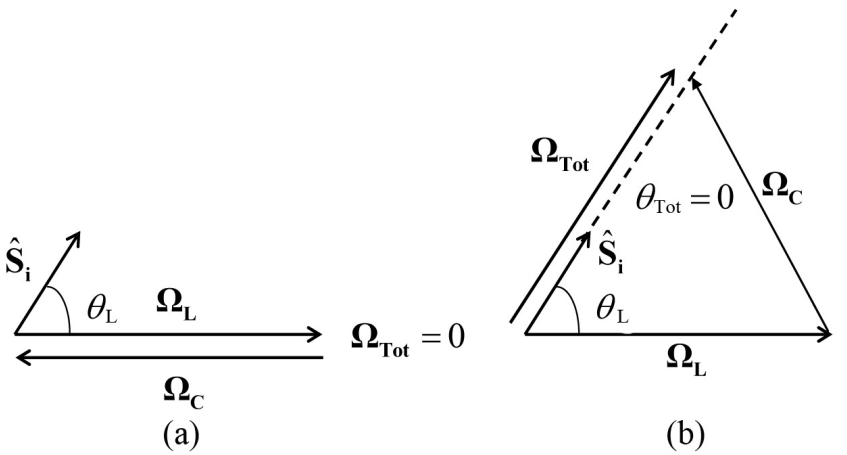


Figure 10.14 Two points of operation of a first-order PMD compensator: (a) $\boldsymbol{\Omega}_{\mathbf{Tot}} = 0$ and (b) $\boldsymbol{\Omega}_{\mathbf{Tot}}$ is parallel to the input polarization state $\hat{\mathbf{S}}_{\mathbf{i}}$.

(i) $\boldsymbol{\Omega}_{\mathbf{Tot}} = 0$: It is easy to understand that $\boldsymbol{\Omega}_{\mathbf{Tot}}$ will be zero if $\boldsymbol{\Omega}_{\mathbf{C}} = -\boldsymbol{\Omega}_{\mathbf{L}}$. However, since $\boldsymbol{\Omega}_{\mathbf{L}}$ varies randomly with time and $\boldsymbol{\Omega}_{\mathbf{C}}$ is fixed, this point of operation cannot be achieved perfectly.

(ii) $\theta_{\mathbf{Tot}} = 0$: This point of operation when $\boldsymbol{\Omega}_{\mathbf{Tot}}$ is parallel to $\hat{\mathbf{S}}_{\mathbf{i}}$ can be rigorously achieved by selecting the value of $|\boldsymbol{\Omega}_{\mathbf{C}}|$ sufficiently large that $|\boldsymbol{\Omega}_{\mathbf{C}}| \geq |\boldsymbol{\Omega}_{\mathbf{L}}|$. It is easy to understand that in such a case, the resultant PMD vector $\boldsymbol{\Omega}_{\mathbf{Tot}}$ can always be made parallel to $\hat{\mathbf{S}}_{\mathbf{i}}$ so that $\theta_{\mathbf{Tot}} = 0$.

The control algorithm is a crucial component of a PMD compensator, as it keeps tracking the local maxima in the signal quality by sending the feedback signal to the polarization controller. In the literature, various qualities of the output signal are used for generating the feedback signal; among these studies,

the DOP is most commonly used. It may be mentioned that all of the time or frequency components of the input pulse have the same polarization states, and hence, the DOP of the input pulse will be 1. However, because the input pulse splits into two pulses that travel with different group velocities, the various components of the pulse at the output end will be different. For example, the SOP of the leading edge, middle edge, and trailing edge of the output pulse will be same as the SOP of the fast PSP, input pulse, and slow PSP, respectively. As a result, the DOP of the output pulse will be degraded and will be less than 1. The controlling algorithm attempts to maximize the DOP of the output pulse by making either $\mathbf{\Omega}_{\mathbf{Tot}} = 0$ or $\theta_{\mathbf{Tot}} = 0$, thus minimizing PMD-induced impairments.

References

1. C. D. Poole and R. E. Wagner, "Phenomenological approach to polarization dispersion in long single-mode fibers," *Electron. Lett.* **22**, 1029–1030 (1986).
2. C. D. Poole, "Statistical treatment of polarization dispersion in single-mode fiber," *Opt. Lett*. **13**, 687–689 (1988).
3. A. J. Barlow, J. J. Ramskov-Hansen, and D. N. Payne, "Birefringence and polarization mode-dispersion in spun single-mode fibers," *Appl. Opt.* **20**, 2962–2968 (1981).
4. A. J. Barlow, D. N. Payne, M. R. Hadley, and R. J. Mansfield, "Production of single-mode fibers with negligible intrinsic birefringence and polarization mode dispersion," *Electron. Lett.* **17**, 725–726 (1981).
5. A. Galtarossa, P. Griggio, A. Pizzinat, and L. Palmieri, "Calculation of mean differential group delay of periodically spun randomly birefringent fibers," *Opt. Lett*. **28**, 1639–1641 (2003).
6. A. Galtarossa, L. Palmieri, and A. Pizzinat, "Optimized spinning design for low PMD fibers: An analytical approach," *J. Lightw. Technol*. **19**, 1502–1512 (2001).
7. C. D. Poole, J. H. Winters, and J. A. Nagel, "Dynamical equation for polarization dispersion," *IEEE Photon. Technol. Lett.* **3**, 68–70 (1991).
8. G. J. Foschini and C. D. Poole, "Statistical theory of polarization dispersion in single-mode fibers," *J. Lightw. Technol*. **9**, 1439–1456 (1991).
9. D. Gupta, A. Kumar, and K. Thyagarajan, "Polarization mode dispersion in single-mode optical fibers due to core ellipticity," *Opt. Commun.* **263**, 36–41 (2006).
10. A. O. Dal Frono, A. Paradisi, R. Passy, and J. P. von der Weid, "Experimental and theoretical modeling of polarization mode dispersion in single-mode fibers," *IEEE Photon. Technol. Lett.* **12**, 296–298 (2000).
11. B. L. Heffner, "Automated measurement of polarization mode dispersion using Jones matrix eigenanalysis," *IEEE Photon. Technol. Lett.* **4**, 1066–1069 (1992).

12. A. Galtarossa, and C. R. Menyuk, Eds., *Polarization Mode Dispersion*, Springer, New York (2005).
13. C. D. Poole, "Polarization mode dispersion measurements based on transmission spectra through a polarizer," *J. Lightw. Technol*. **12**, 917–929 (1994).
14. P. A. Williams and C. M. Wang, "Corrections to fixed analyzer measurements of polarization mode dispersion," *J. Lightw. Technol*. **16**, 534–541 (1998).
15. R. Ulrich and A. Simon, "Polarization optics of twisted single-mode fibers," *Appl. Opt.* **18**, 2241–2251 (1979).
16. S. C. Rashleigh, "Origins and control of polarization effects in single-mode fibers," *J. Lightw. Technol.* **1**, 312–331 (1983).
17. R. E. Schuh, J. G. Ellison, L. M. Gleeson, E. S. R. Sikora, A. S. Siddiqui, N. G. Walker, and D. H. O. Bebbington, "Theoretical analysis and measurement of the effect of fiber twist on the polarization OTDR of optical fibers," *Electron. Lett.* **31**, 1772–1773 (1995).
18. M. J. Li and D. A. Nolan, "Fiber spin profile designs for producing fibers with low polarization mode dispersion," *Opt. Lett*. **23**, 1659–1661 (1998).
19. X. Chen, M. J. Li, and D. A. Nolan, "Polarization mode dispersion of spun fibers: An analytical solution," *Opt. Lett*. **27**, 294–296 (2002).
20. N. Gisin, "PMD and PDL," *J. Opt. Fiber Commun. Rep.* **1**, 1–13 (2004).
21. T. Ono, S. Yamajaki, H. Shimizu, and K. Emura, "Polarization control method for suppressing polarization mode dispersion influence in optical transmission systems," *J. Lightw. Technol*. **12**, 891–898 (1994).
22. D. Pennincks, F. Roy, S. Lanne, and J.-P. Thiery, "Statistical study of dynamic polarization mode dispersion compensation based on degree of polarization monitoring," *Microw. Opt. Technol. Lett.* **26**, 41–43 (2000).

Index

Arun Kumar received his M.Sc. and Ph.D. degrees in physics from the Indian Institute of Technology Delhi (IITD), in 1972 and 1976, respectively. Since 1977, he has been on the faculty of the physics department at IITD, where he has been a professor since 1995. He has been a visiting scientist at the Technical University of Hamburg, Germany (1980–1981); the Opto-electronic Group, Strathclyde University, Glasgow, UK (1988); the National Institute of Standards and Technology, Boulder, Colorado (1993–1994); the University of Nice, France (1996); and the University of Jean Monnet, Saint Etienne, France (several times). He has authored/coauthored more than 90 research papers in international journals and has supervised/co-supervised ten Ph.D. theses in the area of fiber and integrated optics.

A perturbation method for the analysis of rectangular-core waveguides, developed by Kumar and coworkers, is now known as the "Kumar method" in the literature. Kumar is a recipient of research fellowships from the Alexander von Humboldt Foundation of Germany (1980–1981) and the Indian National Science Academy (INSA) (1990–1992). His research interests are in the fields of optical waveguides, fiber and integrated optic devices, polarization mode dispersion, and plasmonic waveguides.

Ajoy Ghatak recently retired as a professor of physics from the Indian Institute of Technology Delhi (IITD). He obtained his M.Sc. degree from Delhi University and his Ph.D. degree from Cornell University. Professor Ghatak has held visiting positions in many universities in the United States, Europe, Hong Kong, Australia, and Singapore. His areas of research are fiber optics and quantum mechanics, and he has written several books in these areas. The first edition of his book *Optics* (McGraw-Hill, 2009) has been translated into Chinese and Persian, and his book *Inhomogeneous Optical Waveguides* (Plenum Press, 1977) (coauthored by Professor M. S. Sodha) has been translated into Russian and Chinese. Some of his other books are *Optical Electronics* (Cambridge University Press, 1989) (coauthored by K. Thyagarajan), *Introduction to Fiber Optics* (Cambridge University Press, 1998) (coauthored by K. Thyagarajan), *Quantum Mechanics* (Macmillan India, 2004 and Kluwer, 2004) (coauthored by S. Lokanathan), and *Mathematical Physics* (Macmillan India, 1995) (coauthored by I. C. Goyal and S. J. Chua).

Professor Ghatak is the recipient of several awards, including the 2008 SPIE Educator Award in recognition of "his unparalleled global contributions to the field of fiber optics research, and his tireless dedication to optics education worldwide and throughout the developing world, in particular" and the 2003 Optical Society of America Esther Hoffman Beller Award in recognition of "his

outstanding contributions to optical science and engineering education." He is also a recipient of the CSIR Shanti Swarup Bhatnagar Award, the 16th Khwarizmi International Award, the International Commission for Optics Galileo Galilei Award, and the Meghnad Saha Award (instituted by The University Grants Commission) for outstanding research contributions in theoretical sciences. He received a D.Sc. (Honoris Causa) from the University of Burdwan in 2007.